Министерство образования и науки Российской Федерации
КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

В.В. ТЮРИН, С.Н. ЩЕГЛОВ

ДИСКРИМИНАНТНЫЙ АНАЛИЗ В БИОЛОГИИ

Монография

Краснодар
2015

УДК 573(075.8)
ББК 28.0в7я73
Т 985

Рецензенты:
Доктор сельскохозяйственных наук, профессор,
заслуженный деятель науки Кубани и России
И.А. Драгавцева
Доктор биологических наук, профессор
С.Б. Криворотов

Тюрин, В.В., Щеглов, С.Н.
Т 985 Дискриминантный анализ в биологии: монография / В.В. Тюрин, С.Н. Щеглов. – Краснодар: Кубанский гос. ун-т, 2015. – 126 с. – 500 экз.
ISBN 978-5-8209-1102-6

В монографии рассматриваются теоретические основы и практические аспекты дискриминантного анализа. Используется геометрическая интерпретация существа подхода и характера получаемых результатов, приводятся таблицы основных результатов дискриминантного анализа из программы STATISTICA.

Адресуется студентам и слушателям, проходящим обучение по программам подготовки биологических факультетов университетов, аспирантам, а также научным работникам различных специальностей, интересующимся современным состоянием биологии.

УДК 573(075.8)
ББК 28.0в7я73

ISBN 978-5-8209-1102-6

© Кубанский государственный университет, 2015
© Тюрин В.В., Щеглов С.Н., 2015

Введение

Дискриминантный анализ – метод многомерного статистического анализа – в литературе позиционируется как классификация с обучением. В этом контексте подразумевается одно из предназначений данного анализа для решения биологических проблем – решение задачи отнесения неизвестного объекта к одной из заранее определённых групп. Однако следует отметить, что сфера применения дискриминантного анализа гораздо шире.

В рамках дискриминантного анализа открываются возможности для оценки степени сходства/различия групп путём вычисления расстояний между их центроидами; определения списка признаков из числа учтённых, играющих наибольшую роль в межгрупповых различиях; оценки качества разделения групп.

Основоположником дискриминантного анализа считается Р. Фишер – один из представителей школы английской биометрии и разработчик дисперсионного анализа. Именно он ясно понимал, что остаточная (внутригрупповая) дисперсия, связанная с естественным разнообразием биологических объектов, служит препятствием для объективного сопоставления сравниваемых групп. Единственно возможный способ устранения эффектов внутригрупповой изменчивости – её искусственное апостериорное уменьшение, в идеале сведение к нулю, т.е. минимизация.

Теоретическая база дискриминантного анализа, заложенная Р. Фишером, в дальнейшем получила развитие в трудах Х. Аренса и Ю. Лейтера (1985), А. Афифи и С. Эйзена (1982), М. Кендалла и А. Стьюарта (1976), С.А. Айвазяна с соавторами (1989). Все они подчёркивали, что если наблюдения относятся к нескольким генеральным совокупностям, то для их классификации представляется разумным найти линейную комбинацию признаков, называемую дискриминантной функцией, в которую исходные признаки войдут с некими постоянными величинами – коэффициентами. Задача дискриминантного анализа заключается в определении значений коэффициентов, минимизирующих вероятность ошибочной классификации.

Книга написана биологами и для биологов, поэтому в ней отсутствует строгое математическое обоснование и доказательная база методов дискриминантного анализа и основные его положения представлены в виде уже научно сформулированных постулатов.

1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ДИСКРИМИНАНТНОГО АНАЛИЗА

Задачи дискриминантного анализа

Дискриминация (от лат. *discriminatio*) означает различение, разделение. Именно в различении групп объектов и изучении характера этих различий заключается основная задача дискриминантного анализа.

Такая задача регулярно встаёт практически в каждой области биологии, например: в генетике при сравнении потомств от разных родительских форм на фоне обычно широкой индивидуальной изменчивости особей, в физиологии, биохимии – при сравнении различных серий опытов (или опыта и контроля) на фоне разброса результатов отдельных повторностей; в систематике – при разделении таксонов низкого ранга и определении таксономической принадлежности особи; в популяционной биологии – при сравнении различных популяций вида и т.д.

Дискриминантный анализ относится к категории методов многомерного статистического анализа, предполагающих описание объектов комплексом признаков. Высокая разрешающая способность таких методов обусловлена не только принципом «больше учтённых признаков – больше информации». Не менее важно, что эти методы учитывают систему корреляции признаков.

Как и многие другие многомерные методы, дискриминантный анализ основан на построении линейных комбинаций признаков – функций, в которые каждый из них входит со своим коэффициентом (вкладом). В дискриминантном анализе линейные комбинации называются соответственно дискриминантными функциями:

$$DF = b_1x_1 + ... + b_ix_i + \ldots + b_px_p + C,$$

где DF – значение дискриминантной функции;

x_i – численное значение i-го признака;

b_i – вклад i-го признака в значение функции;

p – число признаков;

C – константа.

Дискриминантный анализ обеспечивает объективное сравнение (разделение) групп за счёт искусственной минимизации внутригруппового разнообразия (дисперсии).

Обратившись к дискриминантному анализу, можно не только оценить достоверность межгрупповых различий и оценить «расстояния» между группами, но и определить те признаки из числа учтённых, которые в первую очередь обусловливают межгрупповые различия. Более того, когда две (или большее число) группы уже разделены в дискриминантном анализе, возможно определить принадлежность неизвестного объекта к одной из них.

Математические основы и геометрическая интерпретация результатов дискриминантного анализа

Исходные данные для дискриминантного анализа должны представлять собой совокупность объектов, разделённых на несколько групп (классов). Число признаков, по которым описаны объекты, не может быть меньше двух и не должно превышать суммарного числа объектов.

Признаки должны быть количественными, а распределение их значений в каждом классе – нормальным. Недопустимо включение в комплекс признаков с единичной (полной) корреляцией.

Описание объекта по комплексу из p признаков геометрически эквивалентно определению его координат в p-мерном пространстве. Единичному объекту в пространстве отвечает точка, группе объектов – «облако точек». Два или более «облака точек», отвечающих изучаемым группам объектов, нередко перекрываются, т.е. неоднозначно различимы. Практически всегда есть особи с такой ординацией в пространстве, которая не позволяет однозначно отнести их к той или иной группе. Задача дискриминантного анализа геометрически формулируется как построение нового пространства, в котором принадлежность объектов к группам определяется однозначно – перекрывание «облаков точек» становится минимальным.

Координаты объектов в новом ортогональном пространстве определяются значениями дискриминантных функций (DF) – специфических линейных комбинаций признаков, определяемых на основе двух корреляционных матриц: межгрупповой (корреляции групповых средних) и внутригрупповой (корреляции признаков особей в группах). Выбор таких осей производится по критерию максимума отношения межгрупповой дисперсии к

внутригрупповой. Математически выбор основан на сравнении матрицы корреляции групповых средних с матрицей корреляции тех же признаков внутри групп.

Суть подхода иллюстрирует рис. 1. Пусть мы имеем две группы объектов, описанных по двум переменным. Эти переменные образуют плоскость, на которой каждая из групп представлена в виде облака точек. Центроиды этих групп обозначены А и В. При рассмотрении проекций крайних объектов каждой группы на координатные оси признаков, например ось X, видны области перекрывания групп. Иными словами, в каждой из групп имеются объекты, положение которых не определено. Они не могут быть однозначно отнесены к той или иной группе. На рис. 1 это объекты в интервале $xb_1 - xa_2$. Но можно построить новую ось (на рис. 1 – DF), значения которой по-иному учтут исходную изменчивость. В результате чего внутригрупповая изменчивость минимизируется и позволяет чётко выразить межгрупповые различия – перекрывание «облаков точек» исключено. Новая ось, т.е. новый интегральный признак, является линейной комбинацией исходных признаков – дискриминантной функцией.

Из рис.1 ясно, что чёткое разделение групп достигается благодаря минимизации внутригрупповых различий. Действительно, расстояния между проекциями объектов a_1, a_2 и b_1, b_2 на ось дискриминантной функции меньше, чем расстояния между их проекциями на ось X. Расстояние между проекциями центроидов групп А и В на ось DF за счёт этого оказывается большим, чем расстояние между проекциями на ось X (и аналогично на ось Y). Это расстояние, называемое расстоянием Махаланобиса, и оценивает действительные различия групп.

На рис. 1 иллюстрируется возможность дискриминантного анализа при разделении двух групп. Но их может быть и больше. Это потребует определения дополнительных дискриминантных функций ортогональных к первой. В общем случае число необходимых дискриминантных функций на единицу меньше числа разделяемых групп. Однако информативная ценность этих функций разная, поскольку они учитывают разную долю исходной изменчивости комплекса признаков. Первая дискриминантная функция учитывает максимум такой изменчивости, вторая – максимум

остатка и т.д. И если первые две-три функции учли значительную (например более 80%) долю исходной дисперсии, то оставшимися функциями можно пренебречь в решении последующих задач дискриминантного анализа.

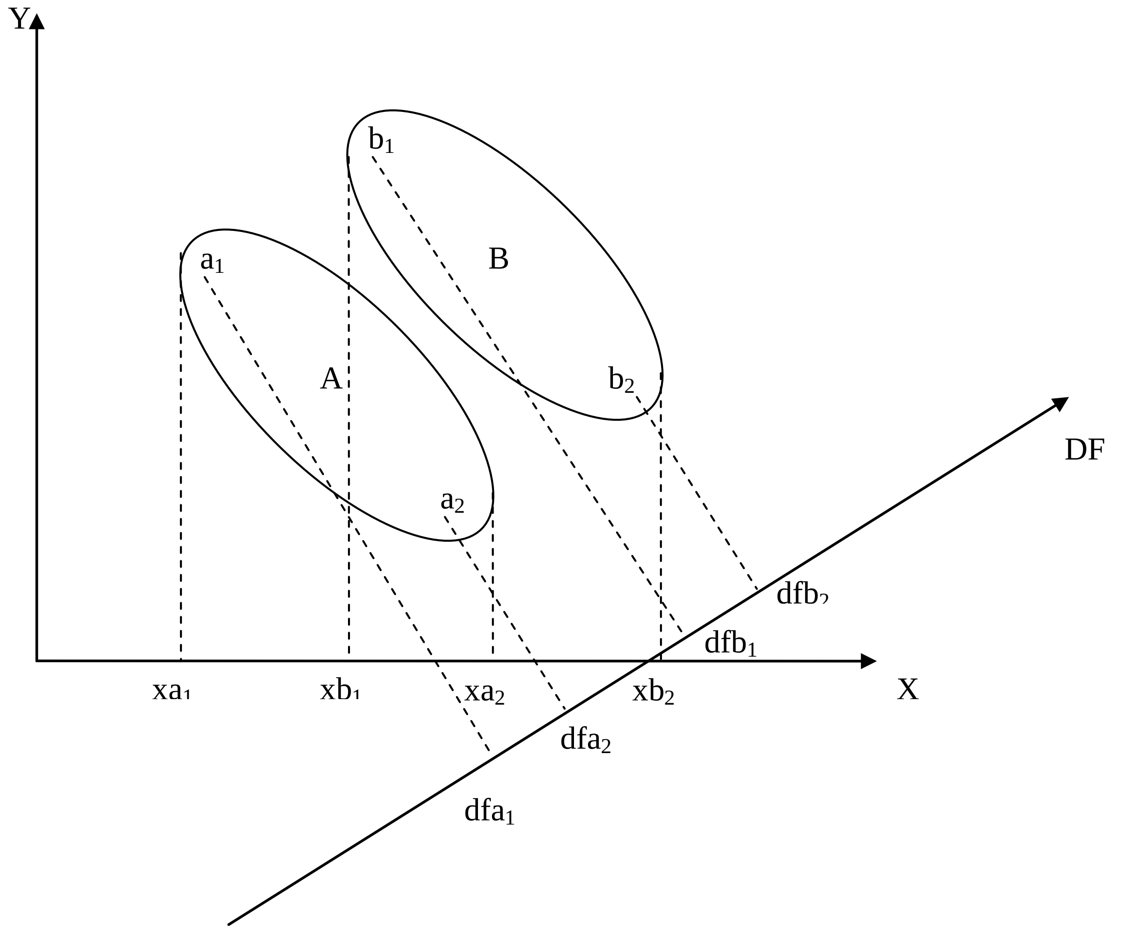


Рис. 1. Геометрическая интерпретация результатов дискриминантного анализа: A, B – центроиды разделяемых групп; xa_1, xb_1 – минимальные значения признака X в группах A и B; xa_2, xb_2 – максимальные значения признака в этих группах; dfa_1, dfa_2 – проекции конкретных объектов группы A на ось дискриминантной функции; dfb_1, dfb_2 – то же для конкретных объектов группы B (по Тюрину В.В., Мореву И.А., Волчкову Ю.А., 2003)

Процент учёта дисперсии дискриминантными функциями не единственный и не самый главный критерий их значимости. Важной характеристикой, позволяющей оценить информативность дискриминантной функции, является коэффициент канонической корреляции (*R*). Каноническая корреляция оценивает меру связи между двумя множествами переменных. Чем больше

значение R, тем выше разделительная способность дискриминантной функции.

Основной критерий оценки эффективности дискриминации для конкретной функции – величина λ-Уилкса. Этот критерий оценивает остаточную дискриминационную способность, под которой понимается способность различать группы, если исключить информацию, полученную с помощью ранее вычисленных функций. Когда остаточная дискриминация мала, вычисление очередной дискриминантной функции не имеет смысла. Статистика λ-Уилкса подчиняется χ^2-распределению и поэтому вычисление данного критерия позволяет оценить достоверность межгрупповых различий для каждой из полученных функций.

По мнению У.Р. Клекка (1989), величины λ-Уилкса близкие к нулю, свидетельствуют о высоком качестве разделения: центроиды групп сильно отличаются друг от друга по отношению к степени разброса наблюдений внутри групп.

В результате дискриминантного анализа в зависимости от числа дискриминантных функций образуется ось, плоскость или пространство, дифференцирующее сравниваемые группы. В любом случае можно оценить степень сходства групп путём вычисления расстояния между их центроидами. В качестве меры сходства, как уже говорилось, выступает расстояние Махаланобиса (D^2), достоверность которого оценивается при помощи критерия Фишера.

Следует отметить, что данное расстояние учитывает корреляционные свойства признаков. Именно этим оно отличается от евклидова расстояния которое оценивает простое сходство групп на плоскости. При переходе в пространство Махаланобиса учитываются все дискриминантные функции, в том числе и малоинформативные. Фактически за счёт этого увеличивается статистический шум и межгрупповые расстояния могут не соответствовать наблюдаемым двумерным распределениям (Ефимов В.М., Ковалёва В.Ю., 2007).

Качество дискриминации групп также может быть проверено с использованием классификационной матрицы. Она вычисляется по классифицирующим функциям, определённым для каждого объекта совокупности. Критерием правильной дискриминации объектов служит процент их отнесения в «свою» группу.

Результат реализации классификационной матрицы можно проиллюстрировать на примере табл. 1.

Таблица 1

Матрица классификаций дискриминатного анализа

Группа	Доля верных отнесений	Число попаданий объектов в группу А	Число попаданий объектов в группу В
А	$c/(c + d)$	c	d
В	$f/(e + f)$	e	f

Примечание. А, В – условные обозначения групп; c – число объектов группы А, отнесённых в область распределения своей группы; d – число объектов групп А, отнесённых в область распределения группы В; e и f – соответствующее число объектов группы В.

Чем выше доля отнесений объектов в «свою» группу, тем лучше качество дискриминации и меньше вероятность ошибок при классификации неизвестных объектов.

Коэффициенты дискриминантной функции

В дискриминантном анализе различают два типа коэффициентов: нестандартизованные и стандартизованные. Посредством коэффициентов первого типа можно получить уравнение для каждого объекта, позволяющее вычислить собственно значение дискриминантной функции. В данное уравнение помимо значений признаков, умноженных на значения коэффициентов, также входит константа. Нормированные коэффициенты приводят к дискриминантным значениям, измеряемым в единицах стандартного квадратического отклонения, т.е. можно определить на сколько сигм конкретный объект отклоняется от главного центроида.

Стандартизованные коэффициенты позволяют судить об относительном вкладе конкретной переменной в дискриминантную функцию. То есть анализ абсолютных величин стандартизованных коэффициентов может дать ответ на вопрос, какова роль каждого признака в межгрупповых различиях. Если абсолютная величина коэффициента для данной переменной у всех значимых дискриминантных функций мала, то эту переменную можно исключить и тем самым сократить число переменных. Эта процедура называется «определение информативного списка признаков».

К числу основных результатов дискриминантного анализа относят также вычисление структурных коэффициентов. Они представляют собой коэффициенты корреляции между конкретными переменными и дискриминантными функциями. Когда абсолютная величина такого коэффициента велика для конкретного признака, принято считать, что вся информация о дискриминантной функции заключена в данной переменной. Это позволяет интерпретировать дискриминантные функции в соответствии с биологической ролью наиболее тесно скоррелированных с ней переменных.

Классифицирующие функции

При решении задачи классификации – отнесении неизвестных объектов в одну из известных групп – применяются классифицирующие функции. Они представляют собой уравнения, составленные для каждой группы. Неизвестный объект относится к классу, у которого в результате решения каждого из уравнений значение классифицирующей функции оказывается наибольшим.

Выбор информативного комплекса признаков

Понятие «информативный комплекс» предполагает, что в анализ будут вовлечены не все признаки, а их некоторая часть, содержащая максимум информации о различии между сравниваемыми группами объектов.

Изначально исследователь руководствуется вполне обоснованным принципом: больше признаков – больше информации об объекте. Однако первоначально избыточный комплекс переменных рационально сократить. Оснований к этому как минимум два:

– исключить малозначимые в разделении групп признаки, составляющие так называемый статистический шум;

– сократить объем последующих измерительных операций с объектом за счёт сокращения первоначального списка признаков.

Арсенал методов дискриминантного анализа содержит два подхода к выбору информативного комплекса признаков, которые мы назвали: алгоритмический и субъективный.

В рамках алгоритмического подхода список информативных переменных формируется автоматически в соответствии с тем или иным алгоритмом пошагового дискриминантного анализа. Предусмотрены два его варианта: «Forward stepwise» и «Backward stepwise» – метод последовательного пополнения и метод последовательного исключения списка информативных признаков. В первом случае выбирается переменная, вносящая наибольший вклад в межгрупповые различия, и далее к ней на каждом шаге анализа присоединятся другие информативные переменные. Метод исключения основан на альтернативной процедуре: из полного списка признаков на каждом шаге анализа последовательно исключаются малозначимые переменные.

Х. Аренс и Ю. Лейтер ввели критерий качества разделения групп – многомерный дистант. По их мнению, окончание процесса формирования информативного списка признаков завершается при достижении компромисса числа признаков и значения дистанта (Аренс Х., Лейтер Ю., 1985).

Субъективный подход основан исключительно на мнении исследователя о роли конкретного признака, формируемого по результатам сопоставления абсолютных значений стандартизованных коэффициентов в значимые дискриминантные функции. Термин «субъективный» означает, что именно исследователь принимает решение, какой стандартизованный коэффициент можно считать большим, а какой – малым.

Результаты исследований, проведённых ранее на основе данных по сопоставлению популяций рыб (Тюрин В.В., Волчков Ю.А., 2009; Тюрин В.В., 2010), свидетельствуют об ожидаемом и важном обстоятельстве. Список признаков, входящий в такой комплекс, отнюдь не является универсальным и изменяется не только от вида к виду, но и в пределах одного вида в зависимости от возраста рыб, условий их выращивания и состава сопоставляемых групп. По первому, но ошибочному впечатлению это может породить сомнения в пользе определения такого списка. Несомненно, что состав информативного комплекса однозначно связан со спецификой объектов и условиями их выращивания. Другими словами, список информативен только в рамках конкретного сравнительного эксперимента. Однако этим не умаляется целесообразность его определения. Этот вывод становится очевид-

ным с позиций традиционного в дискриминантном анализе понятия «обучающая выборка». Выполненные на такой выборке детальные морфометрические исследования, в итоге которых и определяется состав сокращённого комплекса, позволяют многократно снизить затраты труда на морфометрическую характеристику остальных особей и целых групп, вовлечённых в данный конкретный эксперимент. Следует подчеркнуть, что время, затраченное на производство измерений, отражает основные трудозатраты проведения сравнительного эксперимента. Время, потраченное на подготовку электронных файлов данных и последующие статистические процедуры, настолько мало, что им можно пренебречь.

Завершая описание определения информативного комплекса признаков, приведём рассуждения, высказанные в книге С.А. Айвазяна и его соавторов (1989). Их суть сводится к следующему. Естественно считать неинформативными те переменные, у которых не отличаются групповые средние (например, в дисперсионном анализе), т.е. представление об информативности можно составить, исходя из одномерного распределения. Однако в случае многомерных данных это неверно. Включение или, наоборот, исключение переменных из числа информативных может быть существенно скорректировано исходя из информации об их взаимной скоррелированности.

2. РЕАЛИЗАЦИЯ ДИСКРИМИНАНТНОГО АНАЛИЗА В STATISICA

Реализация многомерных методов сложна и требует применения специальных статистических компьютерных программ. В предлагаемом пособии дискриминантный анализ иллюстрируется на примере одной из последних версий программы STATISTICA. В ней статистические процедуры (или группы взаимосвязанных процедур) представлены в пункте меню Statistics, вызов процедуры дискриминантного анализа осуществляется в разделах Multivariate Exploratory Techniques – Discriminant Analysis.

Ввод данных может быть осуществлен как при помощи собственных средств программы STATISTICA, так и с использованием других программ (Excel, Statgraphics и др.) с последующим импортом в STATISTICA. Матрица данных должна быть следующего вида:

$$x_{11}\ x_{12}\ \dots\ x_{1p}\ y_1,$$
$$x_{21}\ x_{22}\ \dots\ x_{2p}\ y_2,$$
$$\dots\ \dots\ \dots\ \dots\ \dots$$
$$x_{n1}\ x_{n2}\ \dots\ x_{np}\ y_n.$$

x_{i1} ... x_{ip} представляют собой значения признаков для i-го объекта, n – число объектов, p – число признаков. Переменная y указывает на принадлежность объектов к одному из классов исходной совокупности, подлежащих разделению. При выборе переменных к анализу именно она указывается в поле «Grouping variable».

Рассмотрим реализацию методов линейного дискриминантного анализа в программе STATISTICA на примере изучения различий трёх семей белого толстолобика, описанных по комплексу морфометрических признаков. Семьи поименованы буквами латинского алфавита и цифрами, где буква – имя самки, цифра – имя самца.

Эти данные, на наш взгляд, наиболее удачны для обсуждения возможностей изучения межгрупповых различий в рамках дискриминантного анализа. Во-первых, семьи – потомства от индивидуальных скрещиваний являются группами рыб с гарантированными генотипическими различиями. Во-вторых, мор-

фометрические признаки, измеренные согласно методике И.Ф. Правдина (1966), служат оптимальным способом описания фенотипа объекта.

Обращение к морфологическим признакам для оценки генетического сходства семей (потомств от индивидуальных скрещиваний) растительноядных рыб (Волчков Ю.А., 1994; Тюрин В.В., 2010) было обусловлено не только большей их практичностью или недостаточной изученностью других признаков рыб. Успешное применение морфометрического анализа в решении разных биологических проблем убедило в том, что «морфологические признаки заключают в себе богатейшую информацию об организме в целом» (Заренков Н.А., 1976).

Не вдаваясь в подробное обсуждение причин их высокой информативности, следует отметить только одну, существенную с позиций системного анализа изменчивости. Возможность исследовать одновременно большое число признаков, неформально объединяемых в комплекс, обеспечивает требуемую репрезентативность набора характеристик фенотипа (Животовский Л.А., 1987) и, как следствие, высокое качество его описания. Напомним, что в результате анализа изменчивости именно морфологических признаков был развит ряд важных положений. Так, Ю.А. Филипченко (1929, 1979) обосновал тезис о связи системы фенотипических корреляций признаков с генетической структурой популяции. Р.Л. Берг (1964) получила доказательства того, что образование классов сбалансированного полиморфизма сопровождается возникновением комплексов взаимообусловленных признаков, различных у разных форм одного и того же вида. А.И. Купцов (1948) зафиксировал изменение корреляционной структуры морфологических признаков в результате отбора по продуктивности. А.С. Мержаниан (1928) предложил метод идентификации растений винограда с генетически обусловленными отклонениями в продуктивности и признаках качества, основанный на оценке комплекса морфологических признаков листа.

Исходные данные для реализации процедур дискриминантного анализа представлены в прил. 1. Задача заключается в изучении межгрупповых (межсемейных) различий по комплексу из 22 морфометрических признаков.

Сами данные необходимо ввести либо в программе Excel с последующим их импортом в пакет STATISTICA, либо с использованием собственного редактора данных пакета.

Доступ к модулю «Дискриминантный анализ» осуществляется согласно рис. 2: пункт меню «Statistics», закладка «Multivariate Exploratory Techniques», процедура «Discriminant Analysis».

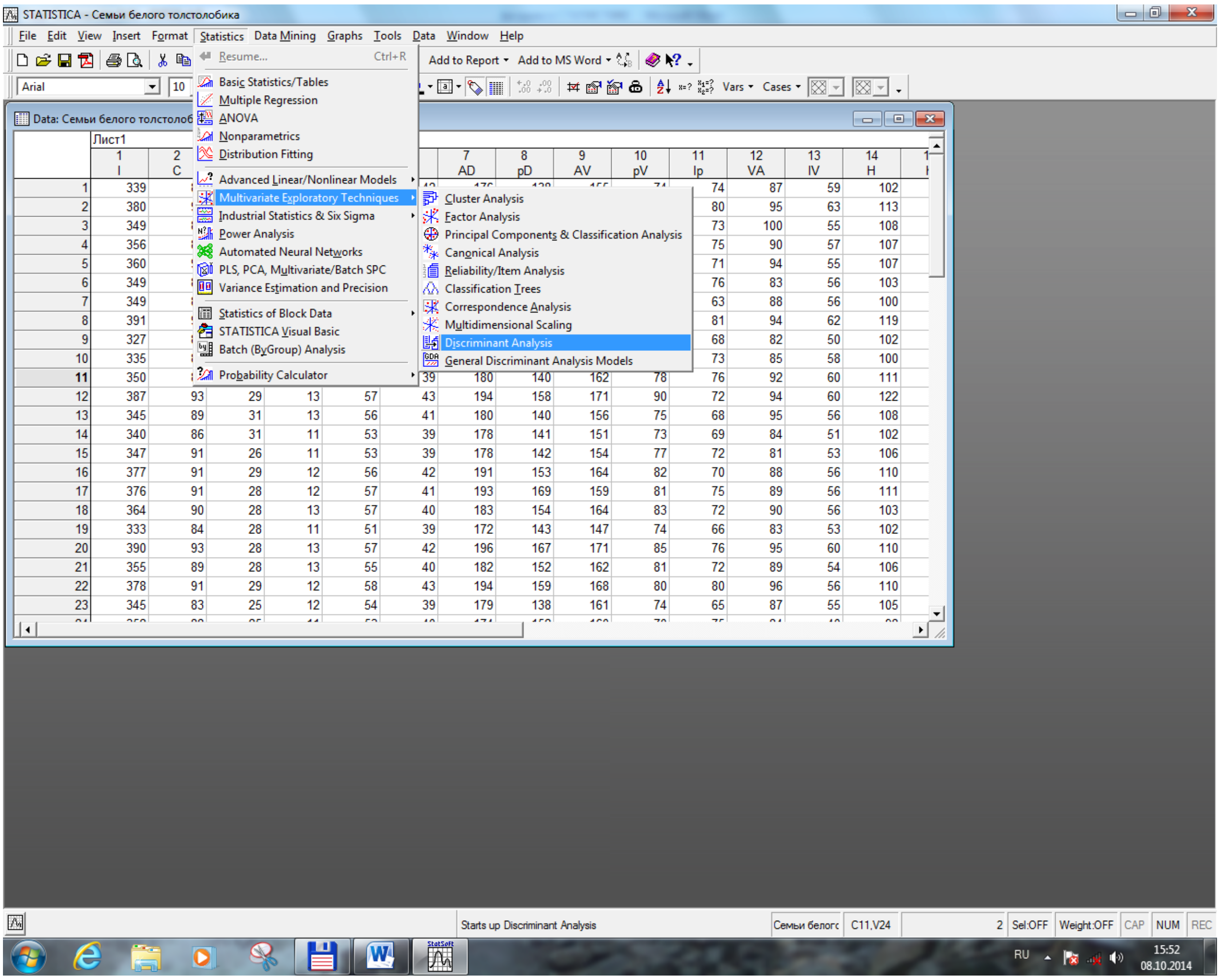


Рис. 2. Выбор модуля «Дискриминантный анализ»

Стартовая панель дискриминантного анализа имеет вид, представленный на рис. 3. В случае маркирования позиции «Advanced options (stepwise analysis)» будет запущена процедура пошагового дискриминантного анализа с автоматическим выбором информативных переменных. Вначале выполним дискриминантный анализ в стандартном режиме.

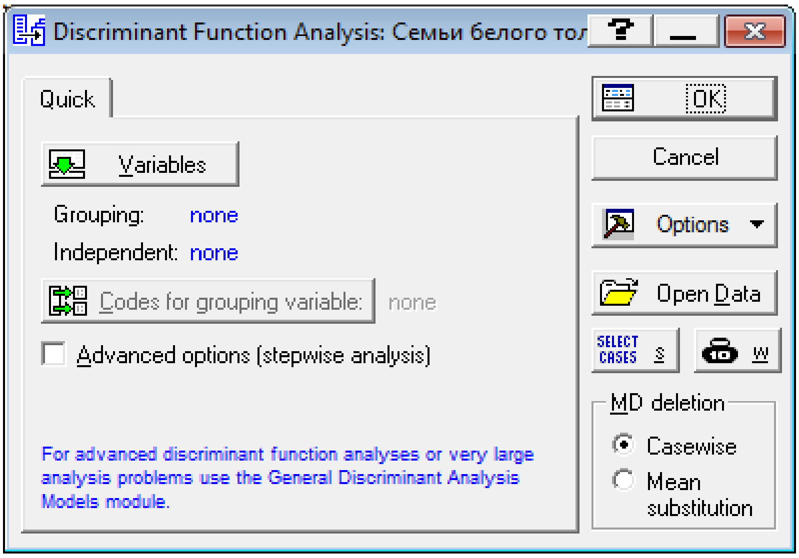


Рис. 3. Стартовая панель модуля «Дискриминантный анализ»

Любая вычислительная процедура в программе STATISTICA начинается с выбора переменных для анализа (рис. 4). После нажатия на кнопку «Variables» откроется панель с двумя окнами: «Grouping variable» и «Independent variable list». В первом окне необходимо указать переменную, идентифицирующую групповую принадлежность объектов. В нашем случае это переменная «Семья», во втором окне выбираем список морфометрических признаков, по которым непосредственно будет проводиться дискриминантный анализ.

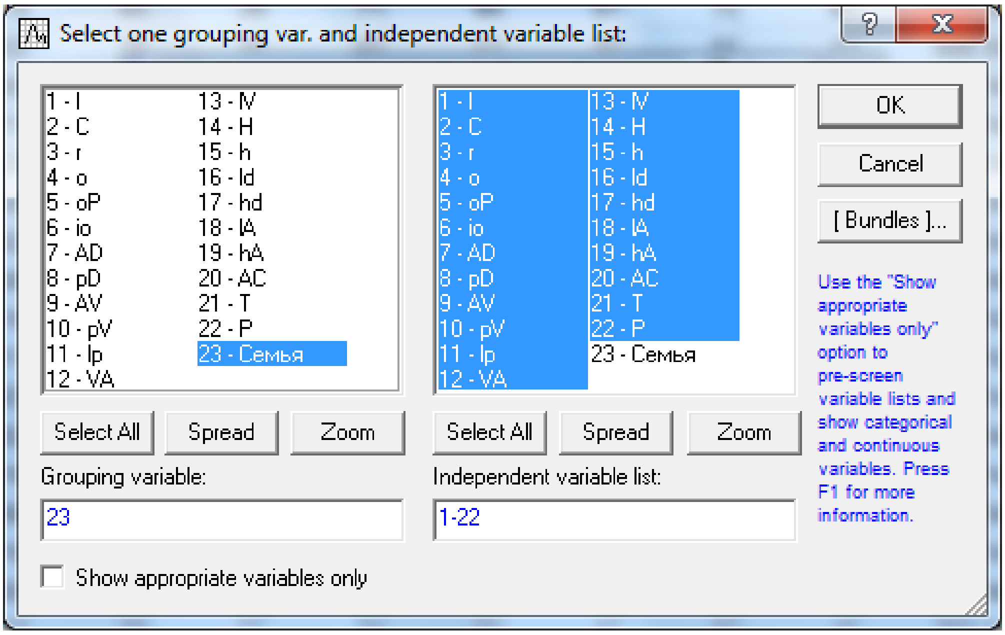


Рис 4. Окно выбора переменных для анализа

Первая панель результатов имеет вид, представленный на рис. 5.

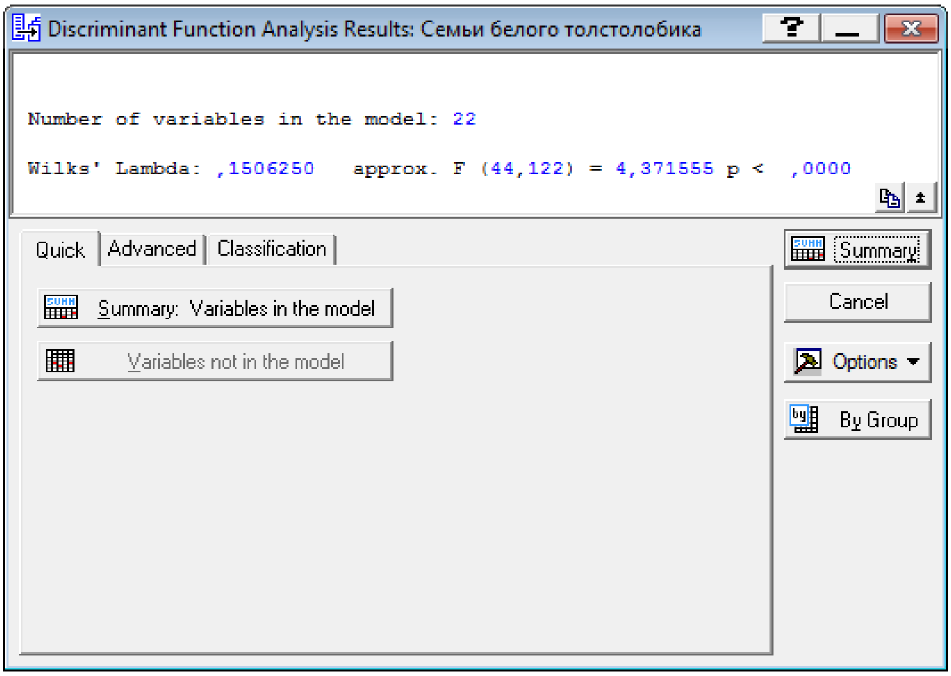


Рис. 5. Панель результатов дискриминантного анализа

Для доступа к основным результатам анализа следует переключиться на закладку «Advanced» (рис. 6).

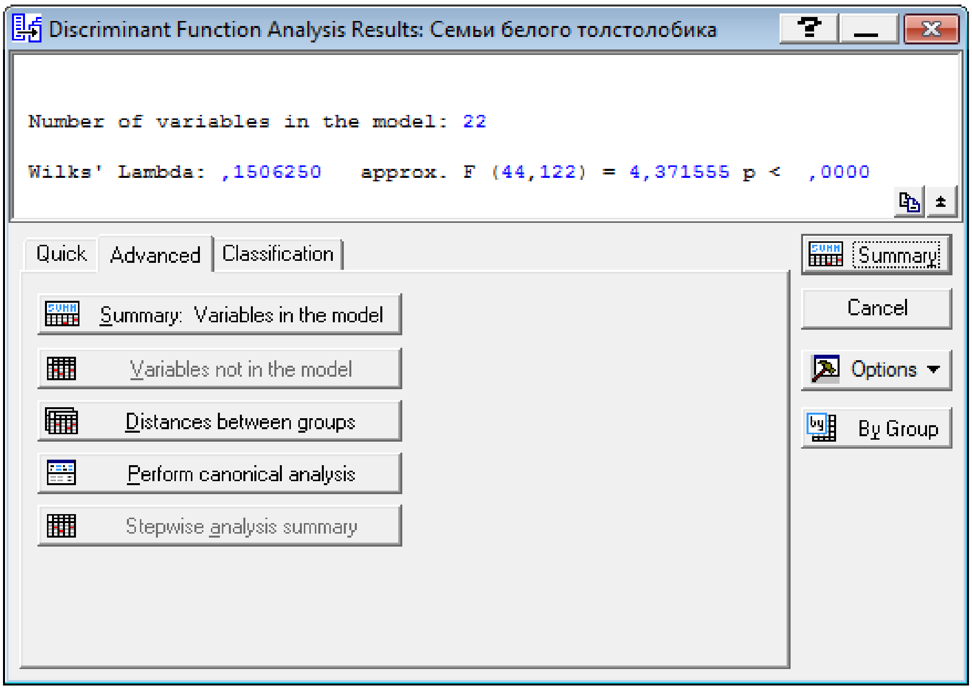


Рис. 6. Меню закладки «Advanced»

В данном меню открывается доступ к таким важнейшим результатам дискриминантного анализа, как «расстояния между группами» (Distances between groups) и «Канонический анализ» (Perform canonical analysis).

После нажатия на кнопку «Distances between groups» на экран выводятся три таблицы: p-levels, F-values, Squared Mahalanobis Distances (рис. 7–9).

Data: p-levels (Семьи белого ...

	p-levels (Семьи белого толстолобика)		
Семья	M3	N2	M2
M3		0,000000	0,000000
N2	0,000000		0,004238
M2	0,000000	0,004238	

Рис. 7. Значения уровней значимости

Data: F-values; df = 22,61 (Семьи белого то...

	F-values; df = 22,61 (Семьи белого толстолобика)		
Семья	M3	N2	M2
M3		6,818081	5,539180
N2	6,818081		2,370712
M2	5,539180	2,370712	

Рис. 8. Значения критериев Фишера

Data: Squared Mahalanobis Distances (Семьи белого тол...

	Squared Mahalanobis Distances (Семьи белого толстолобика)		
Семья	M3	N2	M2
M3	0,00000	15,71978	10,61285
N2	15,71978	0,00000	5,17287
M2	10,61285	5,17287	0,00000

Рис. 9. Значения квадратов расстояний Махаланобиса

Из рис. 7 следует, что все три сравниваемые группы статистически достоверно различаются в пространстве дискриминантных функций (p<0,01). Несмотря на то что программа STATISTICA выдаёт точное значение вероятности ноль-гипотезы об отсутствии различий, на рис. 8 приводятся эмпирические зна-

чения критериев Фишера. Наибольший интерес представляют расстояния Махаланобиса (рис. 9) для всех парных сравнений. Они позволяют судить о степени различий сравниваемых групп. Так, в нашем случае максимальными различиями (15,7 усл. ед.) характеризуются семьи M3 и N2, которые не имеют общих родителей; далее следует результат сравнения M3 и M2 (10,6 усл. ед.) и наибольшим сходством обладают семьи N2 и M2 (5,2 усл. ед.).

Для получения дальнейших результатов о природе дискриминации следует провести канонический анализ (Халафян А.А., 2008). Панель канонического анализа представлена на рис. 10.

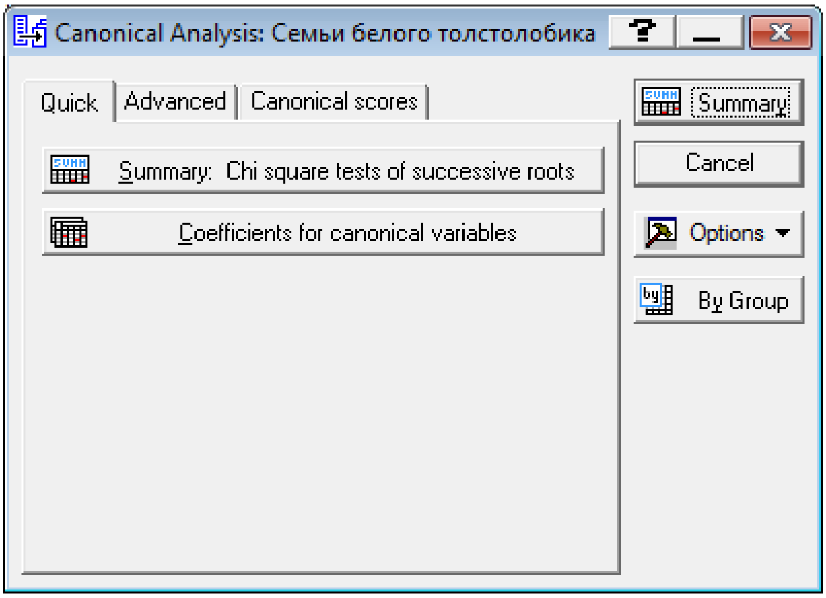


Рис. 10. Панель результатов канонического анализа (закладка «Quick»)

Данная панель имеет три закладки: «Quick», «Advanced» и «Canonical scores». На закладке «Quick» расположены две кнопки «Summary: Chi square test of successive roots» (Итоги: критерий χ^2 последовательности корней) и «Coefficients for canonical variables» (Коэффициенты канонических переменных). После нажатия первой кнопки появится таблица, в которой приведено число построенных дискриминантных функций (корней) и основные статистики, указывающие на их информативность (рис. 11).

Число выделяемых дискриминантных функций должно быть на единицу меньше числа сравниваемых групп, в данном случае две. Обе функции успешно разделяют семьи, о чем свидетельствуют низкие значения уровней значимости (p<0,01).

Data: Chi-Square Tests with Successive Roots Remo...

Roots Removed	Chi-Square Tests with Successive Roots Removed (Сем Eigen-value	Canonicl R	Wilks' Lambda	Chi-Sqr.	df	p-level
0	2,758302	0,856693	0,150625	135,3468	44	0,000000
1	0,766490	0,658715	0,566094	40,6831	21	0,006135

Рис. 11. Основные статистики дискриминантных функций

Вторая кнопка панели выводит на экран две таблицы коэффициентов дискриминантных функций – стандартизованные (Standardized Coefficients) и нестандартизованные (Raw Coefficients). Значения коэффициентов представлены на рис. 12 и 13.

Data: Standardized Coefficients (Семьи белого толсто...

Standardized Coefficients (Семьи белого толстолобика) for Canonical Variables

Variable	Root 1	Root 2
I	-0,30798	-0,31556
C	0,52944	-0,22015
r	0,00799	0,36451
o	0,00301	0,16741
oP	-0,12016	0,00373
io	0,95756	0,20182
AD	1,25639	-1,55076
pD	-0,45292	0,49297
AV	0,22610	0,66306
pV	-0,47562	-0,17731
Ip	-0,11695	0,28780
VA	0,21539	-0,10246
IV	0,08262	0,53523
H	-0,03538	0,74541
h	0,48299	-0,68518
Id	0,24344	0,48131
hd	-0,08325	0,36886
IA	0,66515	0,15948
hA	0,12700	0,10417
AC	0,16639	0,00537
T	-0,32046	-0,27102
P	-2,84606	-0,58324
Eigenval	2,75830	0,76649
Cum.Prop	0,78254	1,00000

Рис. 12. Значения стандартизованных коэффициентов первой и второй дискриминантных функций

В табл. 12 помимо самих значений коэффициентов приведен процент учета исходной дисперсии комплекса признаков (Cum. Prop – накопленный процент). Он дает дополнительную информацию о ценности каждой из функций. В нашем случае первая дискриминантная функция учитывает 78,2% исходной изменчивости, вторая – 21,8%. Если для других данных функций

окажется больше двух, эта информация будет способствовать выбору значимых дискриминантных функций.

Как уже отмечалось, выбор признаков, играющих наибольшую роль в межгрупповых различиях, основан на спосставлении абсолютных значений стандартизованных коэффициентов. Сложность заключается в том, что при выборе надо самостоятельно решить, какой коэффициент следует считать большим, а какой малым по двум функциям одновременно.

Один из возможных подходов для решения данной проблемы может заключаться в суммировании абсолютных значений коэффициентов (табл. 2).

Таблица 2

Абсолютные значения стандартизованных коэффициентов и их сумма

Признак	Абсолютные значения коэффициентов		Сумма коэффициентов
	Root 1	Root 2	
l	0,308	0,316	0,624
C	0,529	0,220	**0,750**
r	0,008	0,365	0,372
o	0,003	0,167	0,170
oP	0,120	0,004	0,124
io	0,958	0,202	**1,159**
AD	1,256	1,551	**2,807**
pD	0,453	0,493	**0,946**
AV	0,226	0,663	**0,889**
pV	0,476	0,177	0,653
lp	0,117	0,288	0,405
VA	0,215	0,102	0,318
lV	0,083	0,535	0,618
H	0,035	0,745	**0,781**
h	0,483	0,685	**1,168**
ld	0,243	0,481	**0,725**
hd	0,083	0,369	0,452
lA	0,665	0,159	**0,825**
hA	0,127	0,104	0,231
AC	0,166	0,005	0,172
T	0,320	0,271	0,591
P	2,846	0,583	**3,429**

П р и м е ч а н и е. Значения сумм коэффициентов информативных признаков выделены жирным шрифтом.

Предположим, что «большим» можно считать сумму коэффициентов больше 0,7. Тогда в информативный комплекс войдут десять признаков: C, io, AD, pD, AV, H, h, ld, lA, P.

Нестандартизованные коэффициенты предназначены для вычисления значений дискриминанатных функций для объектов сравниваемых групп. Это осуществляется путем умножения коэффициентов (с учетом знака) на значения признаков и добавления константы, которая приводится отдельной строкой в таблице коэффициентов. Таким образм, формируется уравнение, решение которого позволяет получить искомое значение дискриминанатной функции.

Data: Raw Coefficients (Семьи белого толст...

Variable	Raw Coefficients (Семьи белого толстопобика) for Canonical Variables	
	Root 1	Root 2
l	-0,0201	-0,02060
C	0,1417	-0,05891
r	0,0033	0,15278
o	0,0044	0,24330
oP	-0,0363	0,00113
io	0,4042	0,08520
AD	0,1552	-0,19162
pD	-0,0615	0,06696
AV	0,0304	0,08914
pV	-0,1106	-0,04122
lp	-0,0261	0,06414
VA	0,0373	-0,01773
lV	0,0152	0,09833
H	-0,0072	0,15248
h	0,2073	-0,29403
ld	0,1026	0,20280
hd	-0,0179	0,07918
lA	0,1651	0,03958
hA	0,0292	0,02399
AC	0,0401	0,00129
T	-0,1224	-0,10352
P	-0,0267	-0,00548
Constant	-32,5977	-4,89604
Eigenval	2,7583	0,76649
Cum.Prop	0,7825	1,00000

Рис. 13. Значения нестандартизованных коэффициентов первой и второй дискриминантных функций

Рассмотрим в качестве примера процедуру вычисления значения первой дискриминантной функции для первого объекта из массива данных (табл. 3).

Таблица 3

Процедура вычисления значения дискриминантной функции по нестандартизованным коэффициентам

Признак	Коэффициент	Значения признака	Произведения
l	–0,0201	339	–6,816300
C	0,1417	88	12,467410
r	0,0033	24	0,080361
o	0,0044	13	0,056844
oP	–0,0363	54	–1,962060
io	0,4042	42	16,978070
AD	0,1552	176	27,323170
pD	–0,0615	138	–8,489230
AV	0,0304	155	4,711207
pV	–0,1106	74	–8,182360
lp	–0,0261	74	–1,928780
VA	0,0373	87	3,243143
lV	0,0152	59	0,895483
H	–0,0072	102	–0,738160
h	0,2073	43	8,912253
ld	0,1026	38	3,897790
hd	–0,0179	59	–1,054410
lA	0,1651	55	9,079478
hA	0,0292	45	1,316155
AC	0,0401	59	2,364379
T	–0,1224	41	–5,018450
P	–0,0267	780	–20,848500
Constant	–32,5977		–32,597700
Сумма произведений и константа			3,689830

В четвертой колонке табл. 3 приведены произведения коэффициентов на значения признаков. После суммирования произведений и добавления константы получим значение 3,689, которое и является искомым значением дискриминантной функции.

Следует оговориться, что данная процедура нами рассматривается лишь в качестве примера реализации алгоритма дискриминантного анализа. Значения дискриминантных функций

для каждого объекта выводятся на экран в соответствующей опции программы STATISTICA, о чем речь пойдёт далее.

Последняя закладка канонического анализа (Canonical scores) позволяет перейти к изучению значений дискриминантных функций объектов выборок (рис. 14).

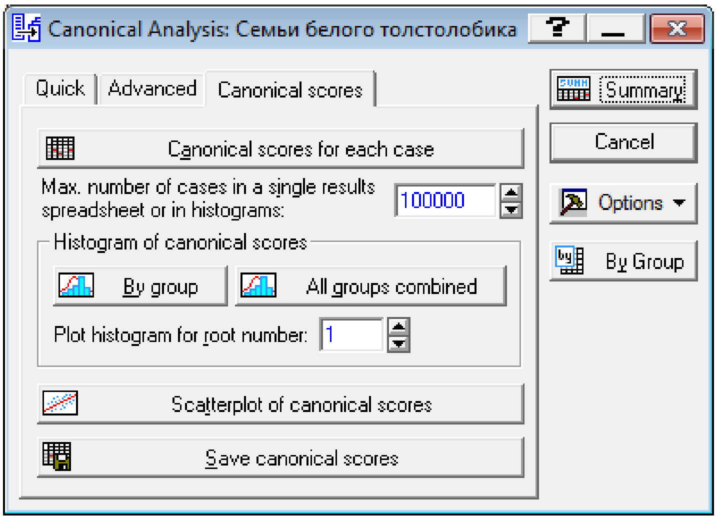
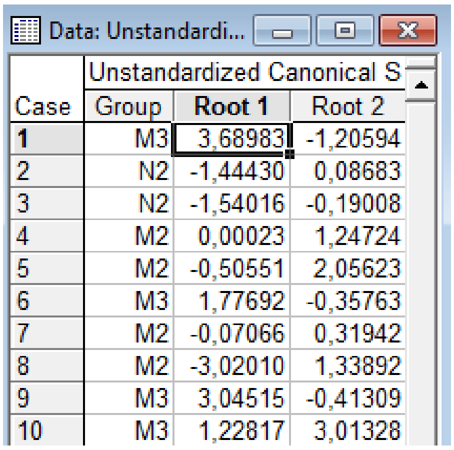


Рис. 14. Панель результатов канонического анализа (закладка «Canonical scores»)

Кнопка «Canonical scores for each case» позволяет просмотреть значения дискриминантных функций для каждого объекта всего массива данных. На рис. 15 в качестве примера приведены значения функций для десяти первых объектов. Курсором отмечено значение первой дискриминантной функции для первого объекта, механизм вычисления которого рассматривался по материалам табл. 2.


Data: Unstandardi...

	Unstandardized Canonical S		
Case	Group	Root 1	Root 2
1	M3	3,68983	-1,20594
2	N2	-1,44430	0,08683
3	N2	-1,54016	-0,19008
4	M2	0,00023	1,24724
5	M2	-0,50551	2,05623
6	M3	1,77692	-0,35763
7	M2	-0,07066	0,31942
8	M2	-3,02010	1,33892
9	M3	3,04515	-0,41309
10	M3	1,22817	3,01328

Рис. 15. Значения первой и второй дискриминантных функций для исходных данных

В области «Histogram of canonical scores» расположены кнопки, позволяющие выводить на экран гистограммы либо для каждой из сравниваемых групп в отдельности, либо для всех групп вместе. Данная процедура, на наш взгляд, представляет интерес в случае сравнения двух групп.

Кнопка «Scatterplot of canonical scores» позволяет построить график двумерного распределения объектов в пространстве дискриминантных функций (рис. 16).

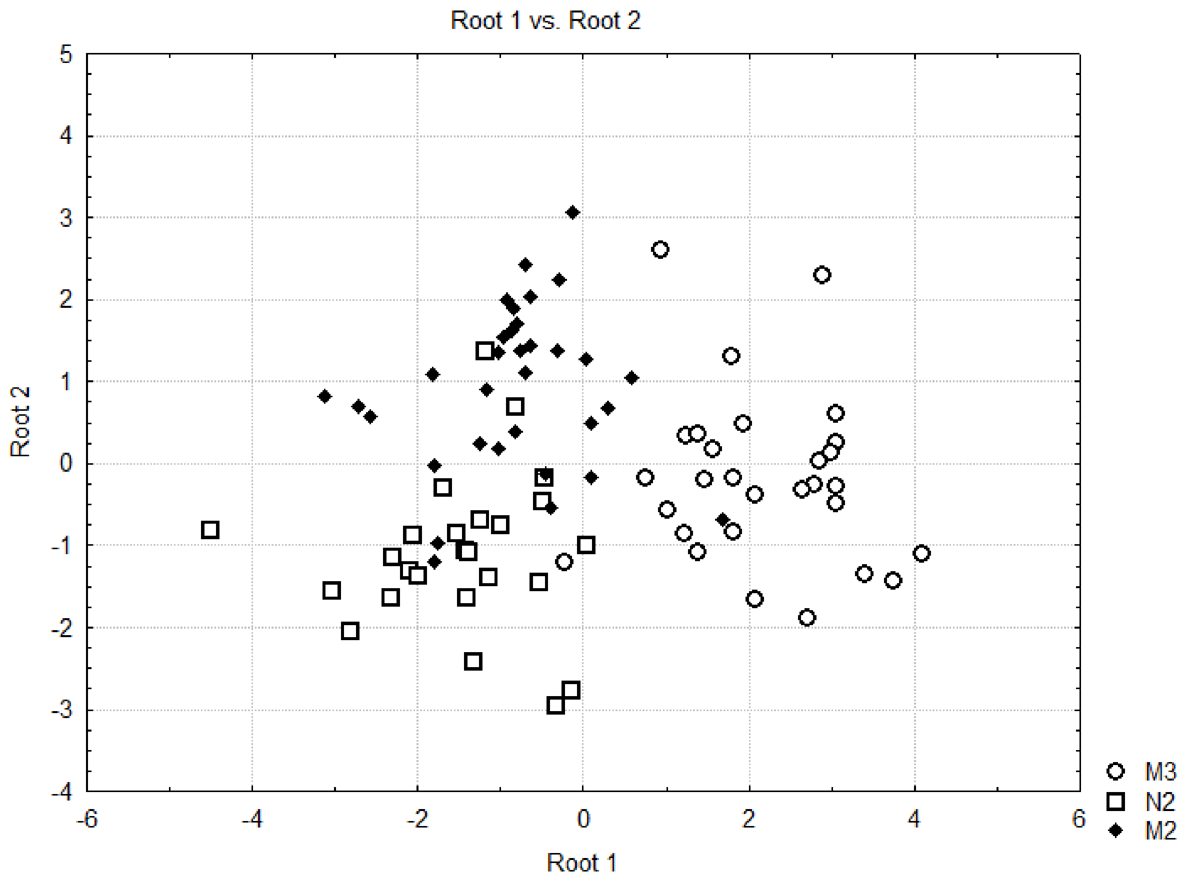


Рис. 16. Распределение выборок трёх семей белого толстолобика в пространстве первой (Root 1) и второй (Root 2) дискриминантных функций

На рис. 16 видно, что «облака точек» каждой из сравниваемых семей располагаются отдельно друг от друга с небольшими областями перегреваний. Так, одна из рыб семьи М2 попадает в область распределения рыб семьи М3, а 4 рыбы семьи N2 – в область распределения семьи М2.

В рамках дискриминантного анализа существует отдельная процедура, позволяющая оценить качество разделения групп. До-

ступ к ней открывается после возврата на первую панель результатов дискриминантного анализа и перехода на закладку «Classification» (рис. 17). Здесь расположена кнопка «Classification matrix», после нажатия которой выводится результат, представленный на рис. 18. Однако следует отметить, что процедура классификации существенно зависит от выбора типа априорных вероятностей «A priori classification probabilities». Предлагаемые варианты выбора вероятностей следующие:

– пропорциональные размерам групп «Proportional to group sizes»;

– равные для всех групп «Sample for all groups»;

– заданные пользователем «User defined».

Последнюю опцию следует использовать в том случае, когда у исследователя есть предварительная информация о возможных различиях в численности сравниваемых выборок, например, известна частота поражённых и здоровых растений, которые мы хотим разделить по комплексу морфологических признаков.

Применительно к нашему примеру следует использовать вариант по умолчанию – «Proportional to group sizes».

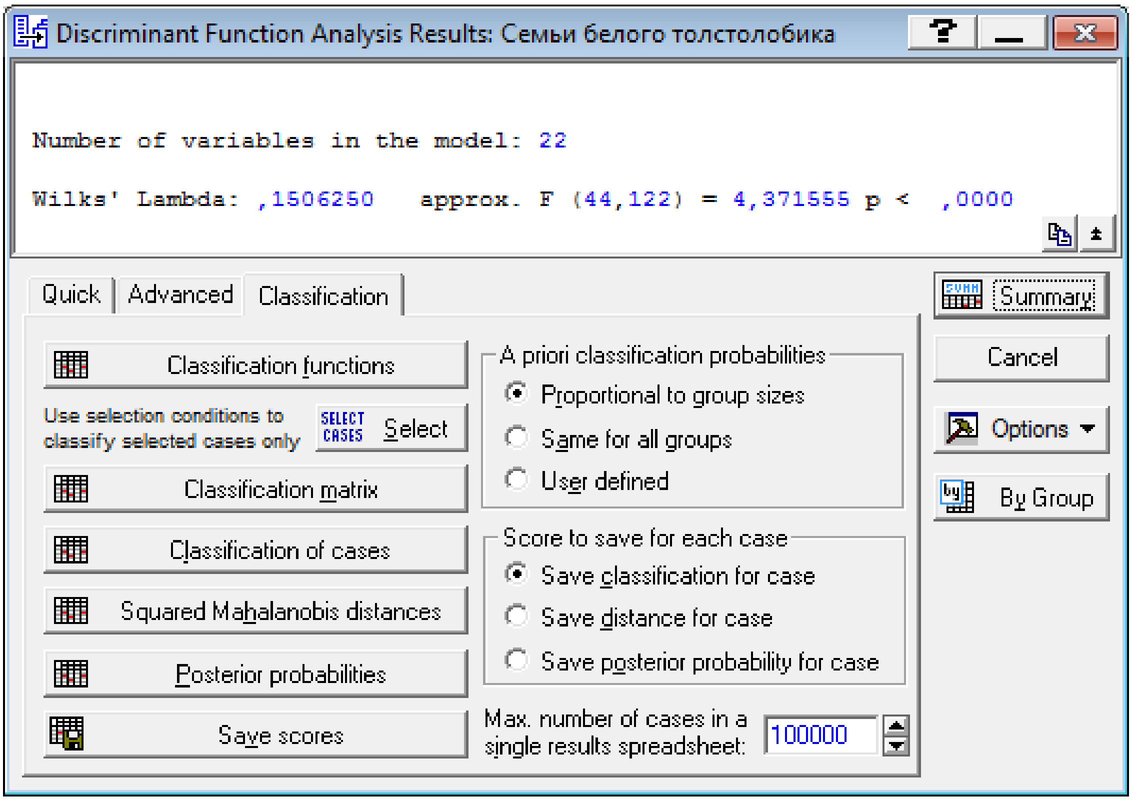


Рис. 17. Содержимое закладки «Classification»

В таблице, приведённой на рис. 18, указывается процент корректных (верных) отнесений рыб конкретных семей в свою область распределения. Так, из 29 рыб семьи М3 27 особей образуют самостоятельное распределение, а две особи попадают в «облака точек» других семей. Точность дискриминации данной семьи рыб составляет 93,1%. Значения априорных вероятностей приведены в первой строке ниже обозначения семей.

Data: Classification Matrix (Семьи...

Classification Matrix (Семьи белого толсто
Rows: Observed classifications
Columns: Predicted classifications

Group	Percent Correct	M3 p=,34118	N2 p=,27059	M2 p=,38824
M3	93,10345	27	1	1
N2	82,60870	0	19	4
M2	84,84849	1	4	28
Total	87,05882	28	24	33

Рис. 18. Классификационная матрица

Data: Classification Functi...

Classification Functions; groupi

Variable	**M3 p=,34118**	N2 p=,27059	M2 p=,38824
I	3,74	3,84	3,78
C	12,22	11,72	11,72
r	-3,59	-3,73	-3,41
o	1,05	0,82	1,34
oP	5,66	5,80	5,77
io	19,58	17,95	18,47
AD	3,90	3,46	3,19
pD	2,06	2,24	2,32
AV	12,32	12,13	12,34
pV	-1,83	-1,37	-1,55
Ip	2,68	2,72	2,84
VA	6,82	6,69	6,68
IV	-0,27	-0,41	-0,19
H	6,44	6,34	6,65
h	20,69	20,14	19,70
Id	15,47	14,89	15,41
hd	3,22	3,22	3,37
IA	9,04	8,37	8,59
hA	3,48	3,34	3,42
AC	3,65	3,49	3,53
T	3,03	3,60	3,27
P	-3,88	-3,78	-3,81
Constant	-3574,26	-3443,61	-3480,45

Рис. 19. Классификационные функции определения семейной принадлежности неизвестных особей белого толстолобика

Данный результат напрямую связан с задачами классификации – отнесения неизвестных объектов к одной из групп, которые в данном случае выполняют роль обучающей выборки. Реализация задач классификации становится возможной при вычислении классификационных функций (Classification functions), которые даны на рис. 19.

Классификационные функции представляют собой уравнения, решение которых основано на перемножении численных значений признаков неизвестного объекта на соответствующие коэффициенты и добавлении константы. Семейная принадлежность объекта определяется максимальным из полученных значений классификационных функций. Более подробно для других объектов данная процедура описана в п. 3.2.1.

Пошаговый дискриминантный анализ

Результаты, представленные в предыдущей главе относятся к реализации дискриминантного анализа в стандартном режиме. Основная цель данного анализа заключается именно в изучении характера различий между группами без задания строгой процедуры выбора информативных переменных. В данном случае помимо реализации первой задачи будет предпринята попытка сокращения числа исходных признаков за счёт устранения тех из них, которые являются излишними и создают, так называемый, статистический шум.

Пошаговый анализ запускается после маркирования позиции «Advanced options (stepwise analysis)» (см. рис. 3) и последующего выбора метода: «Forward stepwise» (пошаговый вперед) или «Backward stepwise» (пошаговый назад). Первый вариант предусматривает последовательное включение значимых переменных в модель, второй – их последовательное исключение из общего списка. Второй вариант реализации дискриминантного анализа является более жёстким, поскольку приводит к максимально возможному сокращению числа информативных переменных.

Рассмотрим вначале пошаговый анализ в режиме «Forward stepwise». После маркирования данной позиции (рис. 20) и нажатия кнопки «OK» будет запущена процедура его реализации.

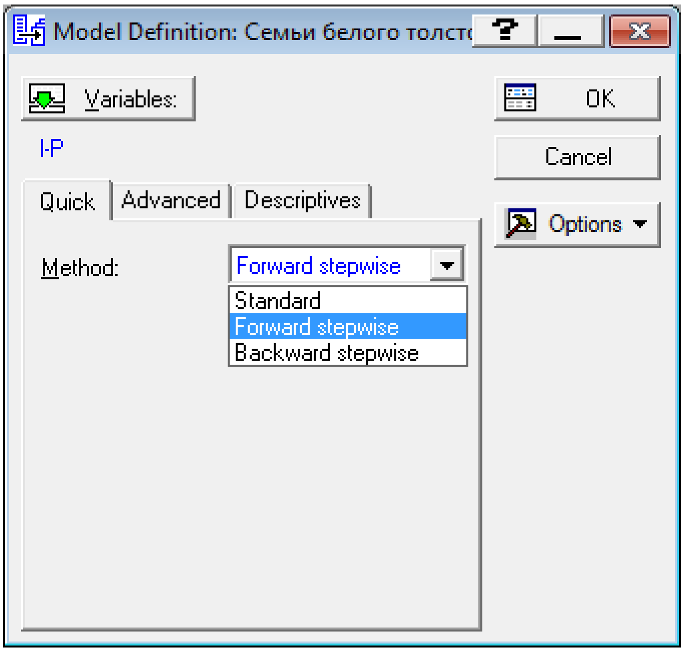


Рис. 20. Выбор дискриминантного анализа в режиме «Forward stepwise»

На закладке «Quick» следует выбрать кнопку «Summary: Variables in the model». После её нажатия будет выведен список переменных, включённых в модель и имеющих наибольшее значение в межгрупповых различиях (рис. 21).

Data: Discriminant Function Analysis Summary (Семьи б...

Discriminant Function Analysis Summary (Семьи белого толст
Step 13, N of vars in model: 13; Grouping: Семья (3 grps)
Wilks' Lambda: ,17030 approx. F (26,140)=7,6635 p< ,0000

N=85	Wilks' Lambda	**Partial Lambda**	F-remove (2,70)	p-level	Toler.	1-Toler. (R-Sqr.)
pD	0,194012	0,877776	4,87350	0,010433	0,473705	0,526295
h	0,216626	0,786142	9,52121	0,000220	0,596305	0,403695
T	0,174029	0,978567	0,76658	0,468461	0,463753	0,536247
IA	0,197177	0,863686	5,52398	0,005922	0,451033	0,548967
P	0,228075	0,746680	11,87419	0,000036	0,061225	0,938775
io	0,218573	0,779139	9,92139	0,000161	0,328498	0,671502
AD	0,213904	0,796145	8,96182	0,000343	0,111750	0,888250
H	0,185431	0,918395	3,10998	0,050819	0,341772	0,658228
IV	0,181140	0,940150	2,22809	0,115320	0,661503	0,338497
pV	0,177148	0,961337	1,40764	0,251560	0,272318	0,727682
Ip	0,179063	0,951058	1,80113	0,172679	0,641254	0,358746
hd	0,177512	0,959366	1,48242	0,234129	0,570771	0,429229
C	0,176036	0,967407	1,17918	0,313565	0,193224	0,806776

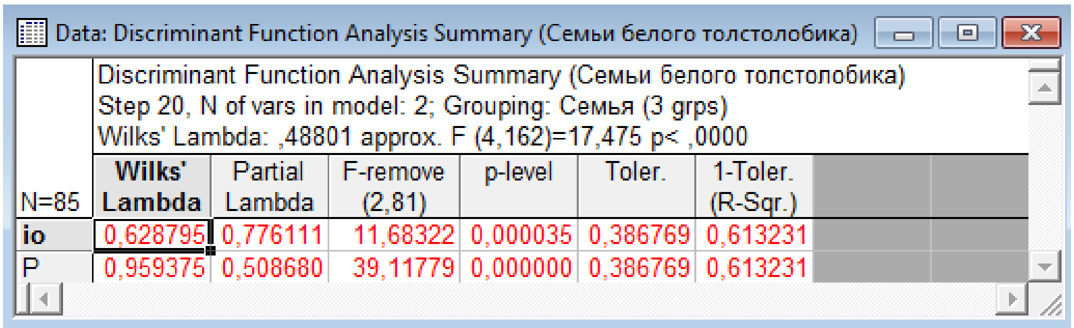
Рис. 21. Результат выбора информативных признаков методом «Forward stepwise»

По результатам пошагового дискриминантного анализа из исходного комплекса 22 морфометрических признаков в модель были включены только шесть: pD – постдорсальное расстояние; h – наименьшая высота тела; lA – длина анального плавника; P – масса тела; io – ширина лба; AD – антедорсальное расстояние.

Сопоставив данный список с тем, который был получен в результате сравнительного анализа стандартизованных коэффициентов, можно обнаружить, что по значениям стандартизованных коэффициентов в качестве информативных было включено большее число признаков – десять. Кроме перечисленных признаков в информативный комплекс вошли: C – длина головы; AV – антевентральное расстояние; H – наибольшая высота тела; ld – длина спинного плавника.

На первый взгляд избыточность числа информативных признаков связана с занижением выбранного минимального уровня суммы стандартизованных коэффициентов. Напомним, что нами был выбран уровень суммы вкладов 0,7 усл. ед. Но более детальный анализ показал, что пошаговый дискриминантный анализ отбирает значимые переменные, ориентируясь не только на их относительный вклад в дискриминантные функции. Примером может служить признак AV. Его суммарный вклад по двум функциям составляет 0,889 усл. ед., что превышает тот же вклад признака lA (0,825). Однако признак AV по результатам пошагового анализа не был включён в модель. Иными словами, его включение не обеспечивало эффективность процедуры разделения групп.

Выполним пошаговый дискриминантный анализ методом «Backward stepwise» (пошаговый назад). Результат анализа представлен на рис. 22.


Data: Discriminant Function Analysis Summary (Семьи белого толстолобика)

Discriminant Function Analysis Summary (Семьи белого толстолобика)
Step 20, N of vars in model: 2; Grouping: Семья (3 grps)
Wilks' Lambda: ,48801 approx. F (4,162)=17,475 p< ,0000

N=85	**Wilks' Lambda**	Partial Lambda	F-remove (2,81)	p-level	Toler.	1-Toler. (R-Sqr.)
io	0,628795	0,776111	11,68322	0,000035	0,386769	0,613231
P	0,959375	0,508680	39,11779	0,000000	0,386769	0,613231

Рис. 22. Результат выбора информативных признаков методом «Backward stepwise»

В данном случае в число информативных признаков вошли только два – io (ширина лба) и P (масса тела).

Возникает вопрос – какова «цена» сокращения исходного комплекса морфометрических признаков, выраженная в качестве разделения групп. Ответ может быть найден в сопоставлении классификационных матриц, полученных на полном и сокращенных списках признаков (рис. 23, 24).

Data: Classification Matrix (Семьи...

Classification Matrix (Семьи белого толст
Rows: Observed classifications
Columns: Predicted classifications

Group	Percent Correct	M3 p=,34118	N2 p=,27059	M2 p=,38824
M3	93,10345	27	1	1
N2	84,00000	0	21	4
M2	87,87878	1	3	29
Total	88,50574	28	25	34

Рис. 23. Классификационная матрица для списка признаков «Forward stepwise»

Data: Classification Matrix (Семьи...

Classification Matrix (Семьи белого толст
Rows: Observed classifications
Columns: Predicted classifications

Group	Percent Correct	M3 p=,34118	N2 p=,27059	M2 p=,38824
M3	82,75862	24	0	5
N2	68,00000	2	17	6
M2	60,60606	6	7	20
Total	70,11494	32	24	31

Рис. 24. Классификационная матрица для списка признаков «Backward stepwise»

Сравнительный анализ процентов корректных отнесений (Percent Correct) позволяет сделать вывод о том, что при сокращении исходного комплекса из 22 признаков до шести методом «Forward stepwise» качество разделения групп не только не ухудшилось, но и увеличилось с 82,6 до 84,0% для семьи N2 и с 84,8 до 87,9% для семьи M2. Этот эффект может быть объяснён снятием статистического шума, выраженного в виде вовлечения в

первоначальный комплекс тех признаков, роль которых в межгрупповых различиях незначительна или полностью отсутствует.

Реализация дискриминантного анализа методом «Backward stepwise» в данном случае приводит к значительному снижению разделения групп. Так, процент отнесения объектов в свою группу для семьи N2 составил только 68,0%, а для семьи M2 – 60,6%. Это означает, что процедура идентификации объектов будет сопровождаться значительными ошибками (32,0 и 39,4% соответственно).

Таким образом, путём сравнительного анализа выборок трёх семей белого толстолобика рассмотрены все основные возможности дискриминантного анализа по изучению межгрупповых различий.

3. ПРАКТИЧЕСКОЕ ИСПОЛЬЗОВАНИЕ ДИСКРИМИНАНТНОГО АНАЛИЗА

В данной главе рассматриваются различные возможности дискриминантного анализа в изучении межгрупповых различий по материалам авторов книги. Некоторые примеры снабжены оригинальными исходными данными. Это позволит читателю самостоятельно реализовать методы дискриминантного анализа для решения конкретных биологических задач.

3.1. Разделение двух групп

3.1.1. Прижизненная идентификация пола у белого толстолобика

Наглядным примером использования дискриминантного анализа может быть задача прижизненной идентификации пола у белого толстолобика в раннем возрасте, когда половой диморфизм ещё чётко не выражен. Актуальность такой задачи вызвана спецификой племенной работы в рыбоводстве – маточные стада должны содержать разную долю самцов и самок. Самцов рационально иметь приблизительно на треть меньше.

Решение поставленной задачи возможно при помощи исследования морфотипа рыб – комплекса промеров тела, включающего в нашем случае около 20 различных измерений по И.Ф. Правдину (1966) (прил. 2).

Исследование проводится в несколько этапов. Вначале из популяции берётся выборка, в которой пол определяется после вскрытия по внешнему виду гонад, и выполняются измерения морфометрических признаков. Такая выборка называется обучающей в том смысле, что именно в результате детального ее исследования можно научиться разделять рыб по полу прижизненно, за счёт измерения определённых морфометрических признаков.

Следующий этап заключается в выполнении дискриминантного анализа на обучающей выборке.

Для разделения двух групп необходима одна дискриминантная функция. Её значения, как и ожидалось, успешно разделили

группы самок и самцов (табл. 4). Однако использовать её уравнение для процедуры идентификации пола представлялось затруднительным. Измерение большого числа (девятнадцати) морфометрических признаков – трудоёмкий процесс, что не позволит провести оценку многочисленного материала.

Таблица 4

χ^2-тест эффективности для выделенных осей
(Chi-Square Tests with Successive Roots Removed)

Roots Removed	Eigenvalue	Canonicl R	Wilks' Lambda	Chi-Sqr.	df	p-value
0	57,08	0,99	0,02	197,00	19	0,00000

Сокращение списка признаков возможно при анализе стандартизованных коэффициентов дискриминантной функции (табл. 5).

Таблица 5

Стандартизованные коэффициенты дискриминантной функции
(Standardized Coefficients for Canonical Variables)

Признаки (Variable)	Дискриминантная функция 1 (Root 1)
L	0,93*
C	–0,24
R	0,26
O	–0,51*
PO	–1,25*
IO	1,17*
HH	–0,10
H	–0,21
AD	0,53*
AV	–0,34
PV	0,51*
LP	–0,32
VA	–0,30
LV	–0,59*
LD	–0,02
HD	0,11
LA	–0,02
HA	0,36
AC	–0,29

В табл. 5 в первом столбце перечисляются условные обозначения морфометрических признаков (L–AC), во втором – значения стандартизованных коэффициентов дискриминантной функции, соответствующие конкретному признаку. Их абсолютные значения варьируют в широких пределах: от 0,02 до 1,25. Среди коэффициентов можно отобрать несколько (в нашем случае семь), которые значительно превышают значения оставшихся. В табл. 5 эти коэффициенты отмечены знаком «*» (обозначение введено авторами). Ранее уже подчёркивалось, что признаки, чьи вклады велики по сравнению с вкладами других признаков, играют более важную роль в межгрупповых различиях. Именно они образуют комплекс информативных признаков. К их числу относятся: L – длина тела рыбы; O – диаметр глаза; PO – заглазничное расстояние; IO – ширина лба; PV – пектовентральное расстояние; LV – длина брюшных плавников.

На сокращённом (информативном) списке признаков дискриминантный анализ был выполнен повторно. Как и в первом случае, единственная дискриминантная функция успешно разделила группы самок и самцов (рис. 25).

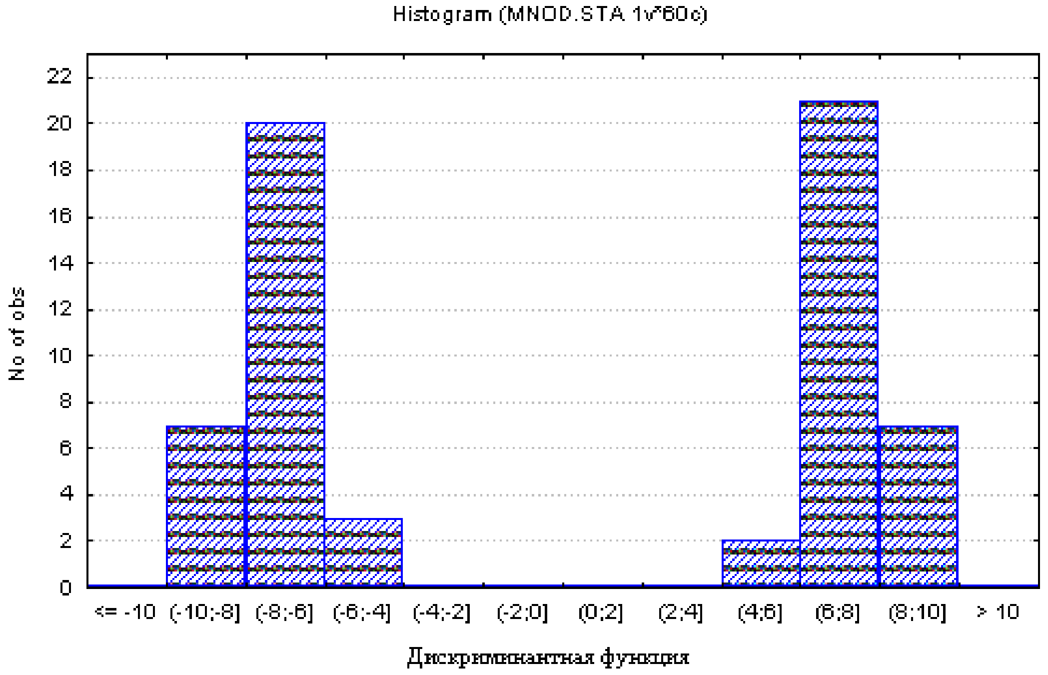


Рис. 25. Распределение групп самок и самцов по значению первой дискриминантной функции

Анализ средних значений искомой функции позволяет определить, что первая диаграмма с областью распределения в отрицательных значениях соответствует самкам, вторая – самцам.

Дальнейшая задача исследований должна заключаться в построении уравнения, позволяющего классифицировать неизвестных особей в группу самок или самцов. Это уравнение можно получить по значениям нестандартизованных коэффициентов дискриминантной функции, полученной на сокращённом (информативном) списке признаков (табл. 6).

Подставив в данное уравнение численные значения признаков неизвестного объекта, можно решить задачу классификации – отнесения его в одну из групп.

Таблица 6

Нестандартизованные коэффициенты дискриминантной функции (Raw Coefficients for Canonical Variables)

Признак	Коэффициент
L	– 0,070
O	0,650
PO	0,830
IO	– 0,860
AD	– 0,006
PV	– 0,060
LV	0,260
Constant	4,660

Проведём классификацию особей, имеющих следующие значения признаков, входящих в информативный комплекс:

L	O	PO	IO	AD	PV	LV
189	9	27	23	95	38	31
195	10	27	21	95	42	30
221	10	24	32	113	46	34
209	10	29	24	108	47	36
216	8	25	30	112	47	32
212	8	22	30	109	42	34
220	9	22	30	114	45	34

Если при умножении соответствующих коэффициентов на значения признаков и прибавления константы получается отрицательное число, то неизвестная особь является самкой, в случае получения положительного значения – самцом.

3.1.2. Идентификация форм устойчивых к коккомикозу представителей рода Cerasus Mill. на основании анализа комплекса биохимических показателей

Одна из главных причин снижения продуктивности насаждений черешни – сильная поражаемость сортов коккомикозом (возбудитель *Coccomices hiemalis Higg*). Развитие болезни на черешне и вишне вызывает сильный листопад, что резко снижает ассимиляционную деятельность растений, эффективность фотосинтеза и ведёт к ослаблению деревьев, снижению урожайности, ухудшению зимостойкости и другим негативным последствиям (Кузнецова А.П. и др., 2012).

Создание устойчивых сортов является наиболее рациональным и безопасным решением проблемы борьбы с болезнью (Кузнецова А.П., Щеглов С.Н., 2009; Щеглов Н.И., 2015).

Данные многолетних наблюдений за поражением в саду грибом черешни и вишни представляют собой наиболее информативный материал для оценки устойчивости образцов, но для ускорения селекционного процесса необходимы экспресс-методы, с помощью которых уже на ранних стадиях развития растения можно дать необходимую оценку, что особенно актуально при возделывании образцов горизонтальной устойчивости.

Материалом для исследований послужили 5 форм рода Cerasus, различающихся по степени устойчивости к заболеванию:

– сорт черешни Франц Иосиф (*C. avium* Moench.), неустойчивый;

– сорт вишни Любская (*C. vulgaril* Mill.), неустойчивый;

– сеянец *Cerasus serrulata* (*C. serrulata* Hally Tolivetta), устойчивый;

– иммунный гибрид №5 (*C. lannesiana* × Франц Иосиф), устойчивый;

– иммунный гибрид №6 (*C. lannesiana* × Франц Иосиф), устойчивый.

В течение двух лет (трёхкратно: весной, летом, осенью) измерялось содержание калия; магния; хлорогеновой кислоты; кофейной кислоты; яблочной кислоты; янтарной кислоты; лимонной кислоты (в мг/дм^3).

Различия между представителями рода *Cerasus* изучались с помощью дискриминантного анализа – метода классификации многомерных наблюдений по одной из нескольких совокупностей (Щеглов С.Н., 2005; Халафян А.А., 2005) (табл. 7).

Таблица 7

Результаты дискриминантного анализа представителей рода *Cerasus*

Дискриминантная функция	Собственное значение	Лямбда Уилкса	χ^2	Степень свободы	Уровень значимости
1	0,890	0,272	107,95	28	0,00
2	0,682	0,514	55,10	18	0,00
3	0,124	0,866	11,91	10	0,29
4	0,026	0,974	2,16	4	0,70

Из табл. 7 видно, что разделение представителей рода *Cerasus* проходит успешно. Наиболее существенную роль в этом разделении играют первые две дискриминантные функции, в пространстве которых был построен график распределения образцов каждого представителя рода *Cerasus* (рис. 26). Следует отметить, что устойчивые формы располагались в области положительных, неустойчивые – в области отрицательных значений функции.

На рис. 26 наблюдается четкое разделение неустойчивых форм (Франц Иосиф и Любская) (выделено овалом) и устойчивых к коккомикозу (*Cerasus cerrulata*, Иммунная 6 и Иммунная 7) по значениям первой дискриминантной функции. Такое распределение форм, различных по устойчивости к коккомикозу, позволяет объединить устойчивые и неустойчивые формы и в дальнейшем работать с этими двумя большими группами.

Для решения задачи идентификации форм устойчивых и неустойчивых к коккомикозу нами был повторно проведён дискриминантный анализ между этими группами (табл. 8).

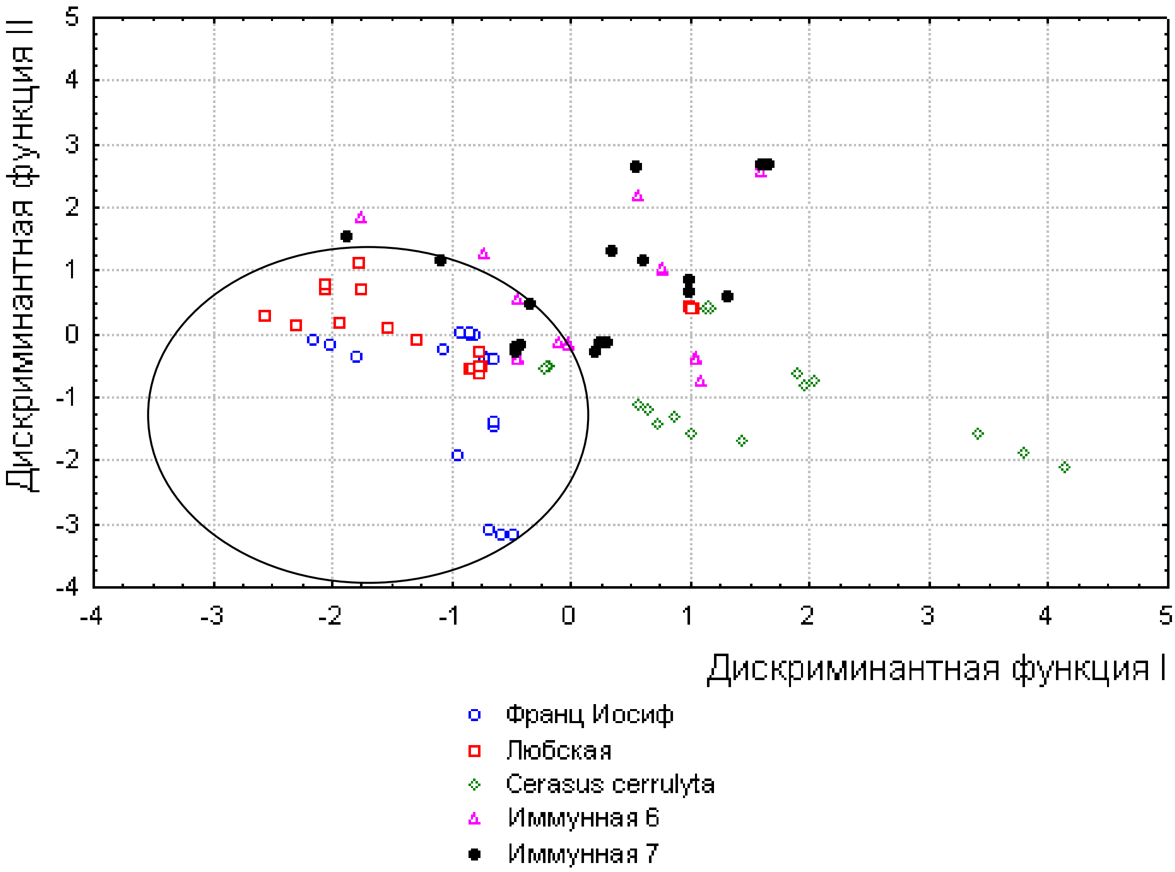


Рис. 26. Распределение образцов представителей рода Cerasus
в пространстве первой и второй дискриминантных функций

Таблица 8

Результаты дискриминантного анализа представителей рода *Cerasus*, устойчивых и неустойчивых к коккомикозу.
χ^2-тест эффективности для выделенных осей
(Chi-Square Tests with Successive Roots Removed)

Roots Removed	Eigenvalue	Canonicl R	Wilks' Lambda	Chi-Sqr.	df	p-value
1	0,833	0,98	0,545	51,22	19	0,00

Как видно из табл. 8, процесс разделения устойчивых и неустойчивых форм прошёл успешно. Были получены уравнения, позволяющие относить неизвестные образцы к устойчивым или неустойчивым к коккомикозу формам:

Устойчивость к коккомикозу = –0,000336 × содержание калия + 0,000578 × содержание магния + 0,000332 × содержание

хлорогеновой кислоты + 0,006650 × содержание кофейной кислоты – 0,126185 × содержание яблочной кислоты + 0,163826 × содержание янтарной кислоты + 0,290296 × содержание лимонной кислоты – 0,359618.

Если в результате подстановки в формулу биохимических показателей нового образца получится положительное число – образец устойчив к коккомикозу, а если отрицательное – неустойчив.

Таким образом, был разработан метод, позволяющий оценивать степень устойчивости неизвестных форм к коккомикозу в раннем возрасте по комплексу биохимических показателей.

3.2. Разделение нескольких групп

3.2.1. Идентификация групповой принадлежности объектов в случае нескольких групп (на примере четырёх пород форели)

В рыбоводстве проблема оперативного различения генетически разных групп по фенотипу стала актуальной по целому ряду обстоятельств. Действительно, решение такой задачи необходимо для организации мониторинга генетической структуры селекционируемого материала (например, в семейной и синтетической селекции), для регистрации новых селекционных достижений, создания и поддержания генетических коллекций, а также разработки научно обоснованных породных стандартов.

Решение этой проблемы требует совершенно определённой экспериментальной модели: наличия нескольких групп (семей, пород, поколений отбора), генетические различия между которыми не вызывают сомнений или весьма вероятны. В решении задачи, сформулированной в названии работы, такая модель выполняет роль обучающей совокупности, исследование которой и обеспечивает оперативное распознавание генетически различных групп (Дерябин В.Е., 1983).

В данном исследовании обучающей совокупностью являлись выборки четырёх пород форели, описанных по комплексу морфометрических признаков.

Ориентация именно на комплекс признаков, формирующих одну систему – морфотип рыбы, уже убедительно обоснована в

большом ряде исследований на различных объектах, в том числе и рыбах. Общий их итог вполне определённо сформулирован в тезисе А.В. Яблокова (1982): переход от анализа изменчивости единичных признаков к их коррелированным комплексам увеличивает генетическую мощность исследования фенотипической изменчивости.

Многочисленные примеры успешного использования системного морфометрического анализа в решении подобных задач по существу не оставляли выбора способа идентификации породной принадлежности форели. В основу должен был быть положен комплекс признаков, составляющих систему «морфотип рыбы». Дальше должен проводиться дискриминантный анализ, минимизирующий внутрипородную дисперсию, с целью наиболее объективной оценки межпородных различий.

Однако реализация этого естественного плана решения задачи натолкнулось на препятствие, заключающееся в существенных различиях массы рыб, как межпородных, так и внутрипородных. Итогом специально проведённого исследования стал вывод о целесообразности исключения массы тела рыб из числа исследуемых признаков и переходу от операций с абсолютными значениями признаков в пользу индексов. Промеры тела следует нормировать на его длину, а промеры головы – на её длину. Это означает, что в основу различения и идентификации генетически различных групп следует положить не комплекс абсолютных значений морфометрических признаков, а комплекс индексов.

Материалом для исследований послужили выборки четырёх пород форели: Камлоопс, Дональдсон, Адлер, Стальноголовый лосось из ремонтно-маточного стада федерального форелеводческого хозяйства «Адлер». Возраст рыб одинаковый. Численность каждой выборки – около 60 особей.

Комплексное морфометрическое описание выборок проводилось по 19 индексам, вычисленным на основе рекомендуемого И.Ф. Правдиным (1966) списка морфометрических признаков.

Статистические методы исследований включали анализ структуры изменчивости с использованием модели однофакторного дисперсионного анализа. Для изучения межпородных различий использовался дискриминантный анализ.

С теоретических позиций структура изменчивости количественных морфометрических признаков и, следовательно, их индексов довольно сложна. В неё входят как собственно генотипические эффекты, так и эффект модификационной изменчивости. Межпородная изменчивость отражает различия групповых генотипов – пород, а внутрипородная – эффекты факторов среды и эффекты внутрипородной генетической гетерогенности.

Однофакторный дисперсионный анализ отдельных индексов показал, что вклад межпородных различий в общую дисперсию варьирует в пределах от 0 до 54%, в массе случаев оставаясь в диапазоне 10–20%. Таким образом, можно заключить, что именно значительная внутрипородная изменчивость служит препятствием для объективной сравнительной оценки пород.

В связи с этим стала ясной необходимость специальных статистических процедур, позволяющих минимизировать внутрипородную изменчивость с целью получения объективной оценки межпородных различий. Такую возможность предоставляет дискриминантный анализ.

В соответствии с числом пород, подлежащих дискриминации, равным четырём, построены три дискриминантные функции, каждая из которых успешно разделяла группы. Но уже первые две из поочерёдно выделяемых функций учли в совокупности 93,9% исходной дисперсии. Иными словами, двумерного пространства достаточно для решения задачи и иллюстрации его результатов.

Качество дискриминации (разделения) пород традиционно оценено по графику ординации особей различных пород в пространстве двух дискриминантных функций, данным классификационной таблицы и матрицы межгрупповых расстояний Махаланобиса (рис. 27; табл. 9).

Из рис. 27 следует, что особи четырёх различных пород действительно эффективно разделяются в пространстве двух дискриминантных функций. Первая из них дифференцирует породы Камлоопс и Адлер, с одной стороны, и форель Дональдсона и Стальноголового лосося – с другой. Вторая разделяет соответственно эти пары: Камлоопс и Стальноголового лосося и Адлер и Долнальдсона. Видное на графике перекрывание «облаков точек»

четырёх пород свидетельствует об определённых проблемах их разделения и заставляет обратиться к строгим количественным оценкам результатов дискриминации. Таковыми являются оценки расстояний Махаланобиса и классификационная таблица дискриминантного анализа.

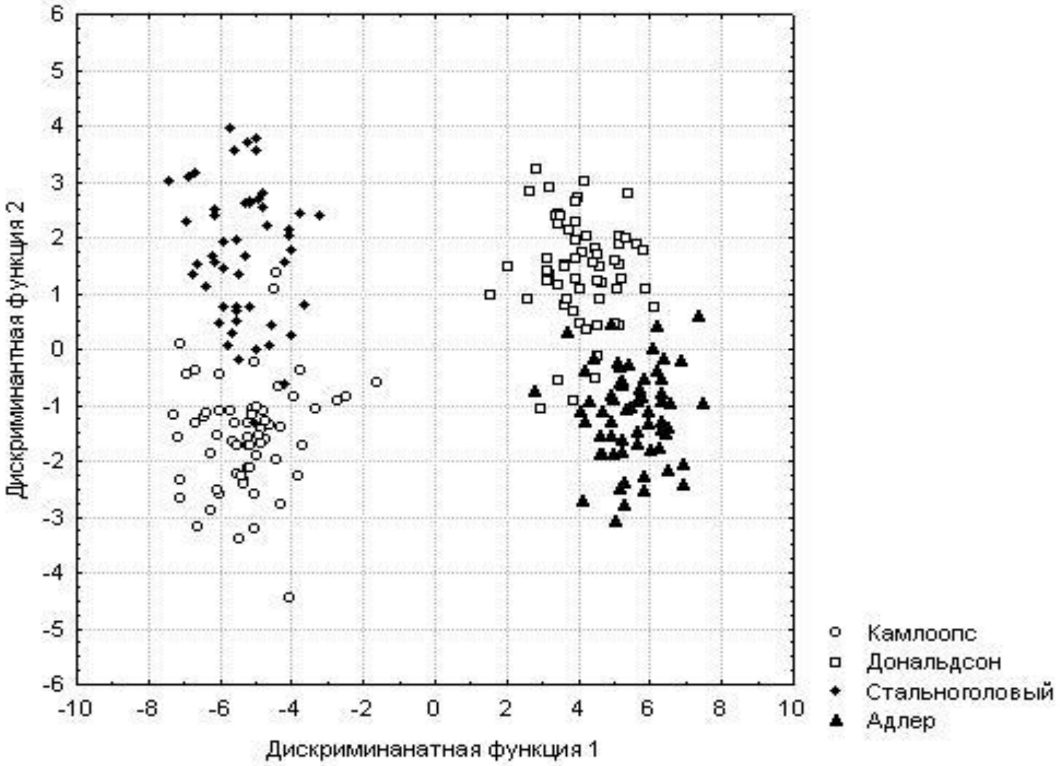


Рис. 27. Ординация особей четырёх пород форели
в пространстве двух дискриминантных функций

Таблица 9

Классификационная матрица пород на полном комплексе морфометрических индексов

Порода	Процент верных отнесений	Число рыб, отнесённых в разделяемые породы			
		1	2	3	4
Камлоопс	95,2	59	0	3	0
Дональдсон	83,3	0	50	1	9
Адлер	88,7	3	4	55	0
Стальноголовый	93,5	0	4	0	58

Матрица убеждает в высокой частоте корректных отнесений, т.е. случаев, когда итоги дискриминантного анализа подтверждают принадлежность особей к «своей» породной группе. Соответствующий показатель (процент верных отнесений) колеблется в пределах от 83,3% для форели Дональдсона до 95,2% для форели Камлоопс.

Таблица 10

Расстояния Махалонобиса между центроидами пород, полученные на полном комплексе индексов

Порода	Камлоопс	Дональдсон	Адлер	Стальноголовый
Камлоопс	–	13,90	8,26	20,05
Дональдсон	0,00	–	12,39	7,20
Адлер	0,00	0,00	–	26,89
Стальноголовый	0,00	0,00	0,00	–

Примечание. Выше главной диагонали приведены межпородные расстояния Махалонобиса, усл. ед.; ниже – вероятность нулевой гипотезы об отсутствии межпородных различий.

Данные табл. 10 окончательно убеждают в эффективности подхода. Все расстояния между центроидами пород в пространстве дискриминатных функций (расстояния Махалонобиса) статистически достоверны.

Дискриминантный анализ пород на основе учёта 20 морфометрических индексов оказывается, безусловно, эффективной процедурой дифференциации пород и идентификации породной принадлежности особей. Но встаёт вопрос о цене решения такой задачи. В теории дискриминантного анализа предусмотрена процедура отбора переменных, наиболее информативных в плане разделения групп. В данном случае она была основана на проведении пошагового дискриминантного анализа с последовательным исключением неинформативных индексов (табл. 11).

В первой колонке табл. 11 приводится список информативных переменных, которые были отобраны для модели, позволяющей наилучшим образом разделить сравниваемые группы.

Таблица 11

Общий анализ дискриминантных функций
(Discriminant Function Analysis Summary)

Индексы	Wilks' Lambda	Partial Lambda	F-remove (3,234)	p-level	Toler.	1–Toler. (R-Sqr.)
pd/l	0,111368	0,875137	11,12894	0,000001	0,916369	0,083631
H/l	0,110921	0,878665	10,77100	0,000001	0,787067	0,212933
av/l	0,122149	0,797897	19,75704	0,000000	0,785219	0,214781
ha/l	0,116931	0,833504	15,58081	0,000000	0,811064	0,188937
h/l	0,136814	0,712374	31,49311	0,000000	0,843695	0,156305
la/l	0,121557	0,801785	19,28289	0,000000	0,940238	0,059762
io/C	0,111665	0,872809	11,36663	0,000001	0,863837	0,136163
r/C	0,121097	0,804832	18,91469	0,000000	0,941079	0,058921
d/C	0,119963	0,812440	18,00706	0,000000	0,918187	0,081813

Примечание. Wilks' Lambda – лямбда Уилкса; Partial Lambda – частная лямбда Уилкса; F-remove – значение F-критерия, связанное с соответствующей частной лямбдой Уилкса; p-level – уровень значимости критерия F-remove; (Toler. – R^2) – коэффициент множественной корреляции данной переменной со всеми другими переменными модели; 1–Toler. (R-Sqr.) – толерантность – мера избыточности переменной в модели.

В группу информативных индексов в плане различения пород форели вошли: отношения к длине тела по Смиту постдорсального расстояния, наибольшей высоты тела, антевентрального расстояния, высоты анального плавника, наименьшей высоты тела, длины основания анального плавника, а также отношения ширины лба, длины рыла, горизонтального размера глаза к длине головы.

В последующих колонках табл. 11 представлены статистические показатели, в той или иной мере характеризующие относительную ценность каждой переменой в межгрупповых различиях. Сюда относятся: лямбда Уилкса, частная лямбда, значения F-критерия, значения вероятности ноль-гипотезы об отсутствии роли переменной и значения толерантности. Ценность той или иной переменной можно определить, в частности, по минимальным значениям частной лямбды и минимальным значениям толерантности ($1 - R^2$) или максимальным значениям собственно $R^2 = 1 -$ Tolerance.

Следует полагать, что выбор информативного (минимально необходимого) в разделении пород комплекса признаков можно считать правильным, если:

– доля общей дисперсии комплекса признаков, учтённой первыми дискриминантными функциями, увеличивается;

– процент корректных отнесений особей к «своей» породе увеличивается или снижается незначительно в сравнении с использованием полного комплекса признаков;

– коэффициент корреляции рангов расстояний Махаланобиса, полученных на полном и сокращённом списке признаков, высокий.

Первое из условий верного выбора информативного комплекса соблюдено. Две первые из трёх возможных дискриминантных функций на полном списке учли 93,9% общей дисперсии, на информативном комплексе – 95,1%. Иными словами, негативный эффект так называемого статистического шума, возникающий за счёт учёта мало значимых для разделения групп признаков, оказался снятым.

Процент корректных отнесений особей к «своей» породе действительно изменился мало. Соответствующие пары значений этого процента для полного и информативного комплекса составили для форели Камлоопс 95,2 и 93,5%; для форели Дональдсона – 83,3 и 80,0%; для форели Адлер – 88,7 и 82,3%; для Стальноголового лосося – 93,5 и 93,6% (табл. 12).

Таблица 12

Классификационная матрица, полученная на информативном комплексе морфометрических индексов

Порода	Процент верных отнесений	Число рыб, отнесённых в разделяемые породы			
		1	2	3	4
Камлоопс	93,5	58	1	3	0
Дональдсон	80,0	0	48	2	10
Адлер	82,3	7	4	51	0
Стальноголовый	93,6	2	2	0	58

И последнее, главное, условие правильности выбора информативного комплекса также оказалось выполненным. Из табл. 13

следует, что ранги парных межпородных расстояний Махаланобиса полностью сохранились, что даже при малом числе парных сравнений позволяет констатировать статистическую достоверность коэффициента корреляции рангов, равному 1,0 на 1%-м уровне значимости.

Таблица 13

Расстояния Махаланобиса между центроидами пород, полученных на полном и информативном комплексах индексов

Порода	Камлоопс	Дональдсон	Адлер	Стальноголовый
Камлоопс	–	11,80 (3)	6,81 (5)	16,44 (2)
Дональдсон	13,90 (3)	–	11,50 (4)	5,12 (6)
Адлер	8,26 (5)	12,39 (4)	–	23,24 (1)
Стальноголовый	20,05 (2)	7,20 (6)	26,89 (1)	–

Примечание. Выше главной диагонали приведены межгрупповые расстояния Махалонобиса, определённые на информативном списке, ниже – на полном списке индексов. В скобках указан ранг расстояния.

Таким образом, можно заключить, что выбранный информативный комплекс из 10 морфометрических характеристик эффективно решает задачу разделения четырёх пород форели.

Итоги проведённой работы позволяют сделать следующие выводы:

– эффект модификационной изменчивости, выраженный в различиях условий выращивания оцениваемого материала и в различиях, накопленных на стадии первого года жизни, требует перехода от абсолютных значений признаков к индексам; такой переход предполагает нормировку промеров головы рыбы на длину головы, промеров тела – на длину тела;

– статистической процедурой, дающей объективную оценку межпородных различий форели, должен являться дискриминантный анализ;

– результаты разделения пород на основе информативного комплекса индексов (9 против 19 изначально учтённых) не только не ухудшаются, но по некоторым показателям улучшаются, что объясняется снятием эффекта статистического шума, выраженного в учёте мало значимых для разделения групп признаков.

Эффективность дискриминантного анализа заключается не только в его способности разделять сравниваемые группы (в данном случае породы) путём построения объективной межгрупповой оценки. Другое его назначение – решение задач классификации, а именно определение групповой принадлежности неизвестных объектов. Математическим инструментом для решения такого рода задач служит построение так называемых классификационных функций. Функции представляют собой уравнения, вычисляемые для каждой группы и состоящие из классификационных коэффициентов и константы. Вектор значений признаков неизвестной особи подставляется в каждое уравнение, что позволяет получить их численные значения. Критерием определения групповой принадлежности неизвестной особи является максимальное из всех полученных значение классификационной функции. Параметры классификационных функций приведены в табл. 14.

Таблица 14

Параметры классификационной функции для идентификации принадлежности особей к одной из четырёх пород форели

Индекс	Коэффициенты признаков и константа функции для каждой из сравниваемых пород			
	Камлоопс	Дональдсон	Адлер	Стальноголовый
pd/l	497,13	467,86	522,75	440,23
H/l	499,46	578,65	487,05	592,05
av/l	–65,19	–15,48	–64,46	34,36
ha/l	145,01	117,15	128,92	109,82
h/l	457,69	474,35	566,19	347,98
la/l	2279,20	2421,87	2374,11	2348,57
io/C	2500,23	2463,01	2396,60	2490,15
r/C	960,58	690,52	911,33	763,47
d/C	1148,26	1092,67	1308,76	1060,26
Constant	–363,88	–1380,36	–1377,05	–1348,37

Нами выбран вариант проверки предлагаемого решения путём идентификации случайно отобранных из выборки самки и самца породы Камлоопс. Векторы значений информативных морфометрических индексов этих особей представлены в табл. 15.

Таблица 15

Значения морфометрических индексов особей, подлежащих породной идентификации

Морфометрический индекс	Значение индексов идентифицируемых особей	
	Самка	Самец
pd/l	0,27	0,28
H/l	0,21	0,20
av/l	0,34	0,32
ha/l	0,72	0,75
h/l	0,24	0,24
la/l	0,09	0,09
io/C	0,37	0,38
r/C	0,48	0,49
d/C	0,10	0,11

В результате арифметических процедур, заключающихся в перемножении значений индексов самки на коэффициенты классификационных функций, были получены следующие значения: 773,4; 630,4; 741,8; 674,30. Видно, что из четырёх полученных значений, соответствующих функциям каждой из четырёх пород, максимально первое. Это означает, что случайно отобранная самка породы Камлоопс совершенно верно отнесена к своей породе. Идентификация самца выполнена столь же однозначно.

Из четырёх значений функций (825,2; 675,6; 793,6; 718,5) простым максимумом обладает первое, соответствующее функции для породы Камлоопс. Можно заключить, что список морфометрических индексов, определённый нами как информативный, не зависит от полового диморфизма и удовлетворительно позволяет решать задачи идентификации породной принадлежности форели.

3.2.2. Сравнительная оценка различных популяций облепихи

На Северном Кавказе облепиха распространена в Краснодарском и Ставропольском краях в пойме р. Кубани, по её притокам Большому Тегеню, Талмыху, Большому и Малому Зеленчуку, Лабе, Белой и Пшишу, Большому и Малому Подкумку, Мару-

хе. В Кабардино-Балкарии, Северной Осетии и Чечено-Ингушетии она растет в долине р. Терека и по его притокам Баксану, Череку, Чечаму, Малке, Нальчику, Ардону, Арыхи, Аргуну, Ассе; в Дагестане – в долине р. Самуре и по его притоку Кухаре, а также по нижнему течению Терека; в Краснодарском крае по р. Лабе, Белой и их притокам (Киртбая Е.К., 1989; Койков Н.Т., 1985).

Целью данного исследования являлась оценка облепихи различных природных популяций. В задачи исследования входило установление сходства и различия популяций по признакам морфологии растения и урожайности.

Растения облепихи были собраны в четырёх эколого-географических зонах: 1) Краснодарском крае, г. Краснодаре; 2) Республике Кабардино-Балкарии, Баксанском ущелье; 3) Чеченской Республике; 4) Краснодарском крае, Лабинском районе, станице Каладжинской.

Изучались следующие признаки: высота растения (м); урожайность (балл); коэффициент шиповатости; количество ягод на 5 см побега (шт.); масса ягод на 5 см (г); длина ягоды (мм); ширина ягоды (мм); длина плодоножки (мм); масса ягоды (г).

Исследование начато с количественной оценки влияния фактора «эколого-географическая зона» на изученные признаки с помощью однофакторного дисперсионного анализа.

Установлено, что эффект условий эколого-географической зоны не оказывает влияния на признаки «масса ягод на 5 см» и «длина плодоножки». Доля соответствующей дисперсии у остальных признаков варьировала от 5,8 (ширина ягоды) до 37,9% (высота растения).

Нами был использован дискриминантный анализ, обеспечивающий объективное сравнение (разделение) групп за счёт искусственной минимизации внутригруппового разнообразия (дисперсии). С помощью дискриминантного анализа представляется возможность выяснить, какие признаки вносят наибольший вклад в различия между популяциями (табл. 16).

Как видно из табл. 16, разделение популяций проходит успешно и статистически значимыми оказываются все три дискриминантные функции.

Распределение растений в пространстве первой и второй дискриминантных функций представлено на рис. 28.

Таблица 16

Дискриминантный анализ популяций облепихи
из разных эколого-географических зон

Дискриминантная функция	Собственное значение	Лямбда Уилкса	χ^2	Степень свободы	Уровень значимости
1	1,66	0,79	148,99	27	0,00
2	0,89	0,68	73,93	16	0,00
3	0,38	0,52	24,84	7	0,00

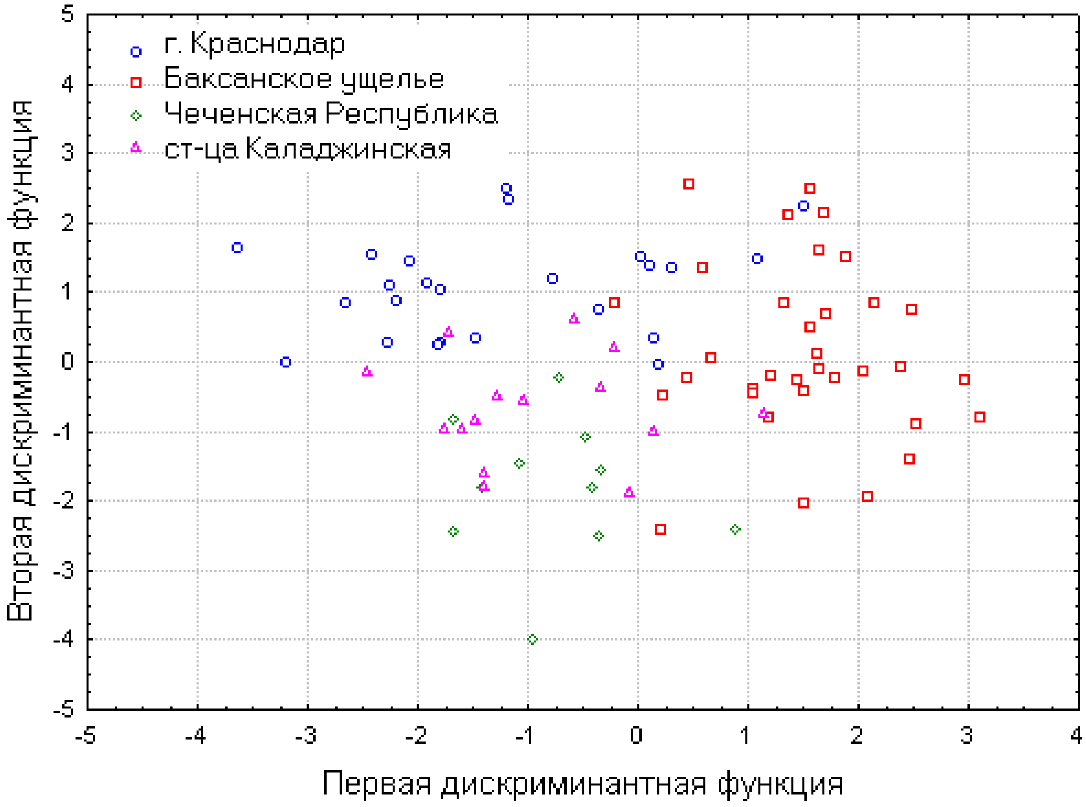


Рис. 28. Распределение растений в пространстве первой и второй дискриминантных функций

На рис. 28 видно, что популяции чётко различаются, хотя некоторые растения оказываются очень сходными с растениями из других популяций.

Стандартизованные коэффициенты (табл. 17) позволяют судить об относительном вкладе конкретной переменной в дискриминантную функцию. Другими словами, анализ абсолютных величин стандартизованных коэффициентов может дать ответ на вопрос, какова роль каждого признака в межгрупповых различиях.

Таблица 17

Стандартизованные коэффициенты дискриминантных функций (Standardized Coefficients for Canonical Variables)

Признак	Дискриминантные функции		
	первая	вторая	третья
Высота растения	–0,194	0,869*	–0,308
Урожайность	1,780*	0,080	–0,268
Коэффициент шиповатости	–0,187	–0,121	0,333
Количество ягод на 5 см	–1,713*	–0,399	–0,527
Масса ягод на 5 см	0,256	0,280	0,332
Длина ягоды	0,158	0,577*	0,184
Ширина ягоды	–0,161	–0,011	0,082
Длина плодоножки	0,160	–0,147	–0,244
Масса ягоды	–0,189	–0,070	0,405

Примечание. * – наиболее весомые вклады в дискриминантные функции.

Нами рассматривались стандартизованные коэффициенты первых двух дискриминатных функций, так как они учли в совокупности 87% общей дисперсии.

Как видно из табл. 17, в разделение популяций наиболее существенные вклады вносят четыре признака из девяти: высота растения, урожайность, количество ягод на 5 см, длина ягоды.

3.3. Определение координат идеального объекта в пространстве дискриминантных функций как метод биологической интерпретации расстояний Махаланобиса

Метод основан на определении параметров идеального объекта – модели, вычислении его координат в пространстве линейных комбинаций и расчете расстояний до него от центроидов групп. Параметры модели должны соответствовать цели эксперимента, максимальное приближение к которой и должно являться критерием отбора. Именно внедрение модели в пространство, разделяющее группы, привносит биологический смысл в безразмерные линейные комбинации (Тюрин В.В., 1994).

Данный подход эффективен в случае, когда из нескольких сравниваемых необходимо выбрать лучшие группы по комплексу показателей. Как правило, это признаки, характеризующие продуктивные качества объектов. Однако метод можно использовать и в иных исследованиях. Так, например, был выполнен выбор оптимального способа обработки семян подсолнечника препаратами биологической и химической природы на основе сопоставления показателей фотосинтетических пигментов (Смирнова Н.С. и др., 2008). В работе В.Г. Назаретяна и др. (2004) по данному принципу был обоснован оптимальный способ лечения дисбиоза у детей с хроническим гастродуоденитом. В качестве параметров модели использовались показатели клинической нормы.

3.3.1. Оценка продуктивности семей белого толстолобика по принципу апостериорной минимизации средовой компоненты изменчивости

Хорошим способом оценки продуктивности групп можно считать анализ результатов эксперимента по определению продуктивности семей белого толстолобика (прил. 3).

Оценка проводилась на втором году жизни в возрасте двухлеток (1+). Выращивание осуществлялось в условиях совместной посадки семей: каждая семья в каждой из четырёх повторностей – прудов (их номера 20, 21, 22, 23). Оценивались три семьи: А2, С2 и С3. Буквой обозначено имя самки, цифрой – номер самца.

Генетически различные группы рыб сопоставляли по комплексу из пяти показателей. Первую их категорию, названную «признаки – компоненты продуктивности», составили средняя конечная масса рыб – mk (г), выход рыбы с единицы площади пруда (рыбопродуктивность) – pr (кг/га) и выживаемость – ft (%). Во вторую, именуемую «параметры рыбоводного эксперимента», вошли средняя посадочная масса рыб – m0 (г) и плотность посадки – pp (шт/га).

Особенность подобных экспериментов в рыбоводстве заключается в отсутствии реальных механизмов по регулированию условий выращивания сравниваемых групп (пород, семей и т.д.). Каждая селекционируемая группа должна быть высажена в отдельный водоем (пруд). При этом пруды могут значительно от-

личаться друг от друга по целому ряду средовых характеристик. К их числу относят: степень развития естественной кормовой базы водоёма, гидрохимический режим, количество рыбы на единицу площади и др.

Зачастую складывается ситуация, когда генотипические различия групп оказываются скрытыми на фоне значительной модификационной изменчивости. Это и было зафиксировано при анализе результатов выращивания каждой из трёх семей в разных прудах, которые в статистическом смысле выступали в качестве повторностей опыта. Однофакторный дисперсионный анализ не выявил достоверных различий между семьями по признакам – компонентам продуктивности. Вся их изменчивость была поглощена случайной, остаточной изменчивостью, обусловленной различиями между прудами. Решение селекционной задачи выбора лучшей семьи по комплексу признаков продуктивности оказалось невозможным в силу высокой внутригрупповой (межпрудовой) изменчивости.

Данные результатов выращивания семей белого толстолобика были проанализированы с использованием дискриминантного анализа. Группы повторностей опыта образованы по семейной принадлежности рыб. В этой ситуации модификационная изменчивость, связанная с различием условий выращивания (разными прудами), должна быть минимизирована с целью выявления и оценки генотипических различий между семьями.

Рассмотрим основные результаты дискриминации. Для удобства работы с программой STATISTICA таблицы результатов будут даваться максимально приближенными к выдаваемым самой программой.

Из табл. 18 видно, что для дискриминации трёх семей построено две дискриминантные функции (0–1). Оценка остаточной дискриминации, выполненная с использованием критерия χ^2, показала, что обе функции являются эффективными в плане разделения семей ($p<0,05$). Это хорошо иллюстрирует график распределения точек опытов по выращиванию семей в пространстве первой и второй дискриминанатных функций (рис. 29).

Из рис. 29 видно, что 12 точек результатов экспериментов (3 семьи в 4 прудах) группируются исключительно по семейной принадлежности.

Таблица 18

χ^2-тест эффективности для выделенных осей
(Chi-Square Tests with Successive Roots Removed)

Roots Removed	Eigenvalue	Canonicl R	Wilks' Lambda	Chi-Sqr.	df	p-value
0	47717,42	0,9999	0,000005	85,0393	10	0,00000
1	2,96	0,6644	0,252737	9,6278	4	0,04718

Примечание. Roots Removed – построенные дискриминантные функции; Eigenvalue – собственное значение; Canonicl R – коэффициент канонической корреляции R; Wilks' Lambda – показатель лямбда Уилкса; Chi-Sqr. – критерий χ^2; df – степень свободы; p-value – вероятность нулевой гипотезы об отсутствии различий.

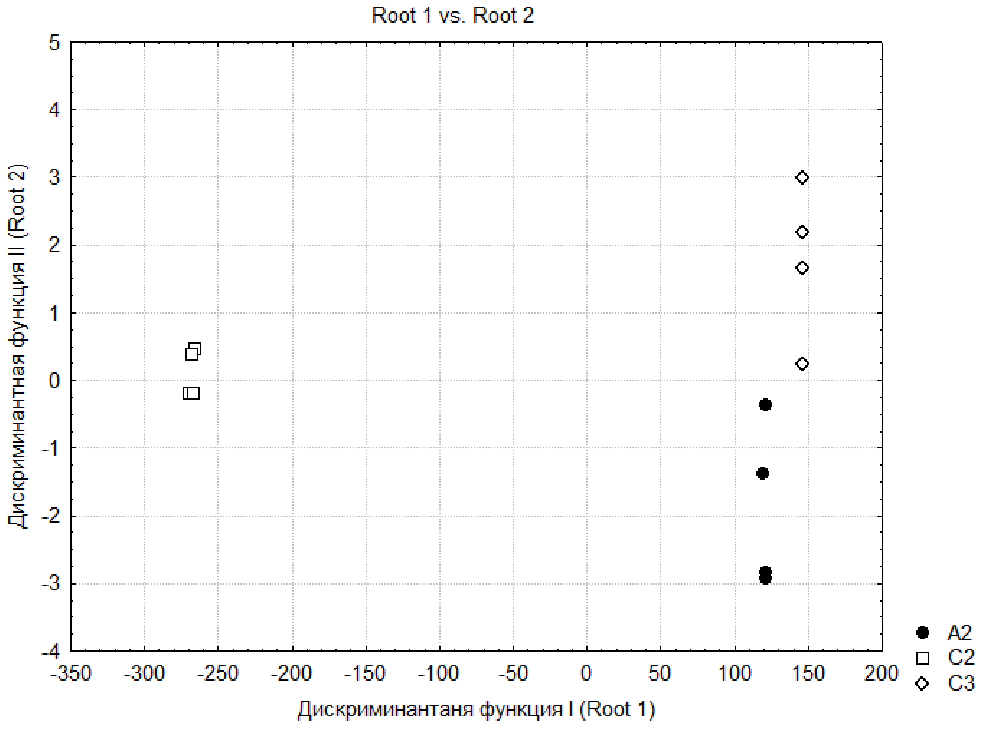


Рис. 29. Распределение точек опытов по выращиванию семей белого толстолобика в пространстве первой и второй дискриминанатных функций (Roots)

Когда задача разделения семей решена – получено дифференцирующее их пространство, можно перейти к процедуре их сравнения путём оценки межсемейных расстояний – расстояний Махаланобиса (табл. 19 и 20).

Таблица 19

Квадрат расстояния Махаланобиса
(Squared Mahalanobis Distances)

Семья	Squared Mahalanobis Distances		
	A2	C2	C3
A2	0,0	150410,7	675,2
C2	150410,7	0,0	171026,7
C3	675,2	171026,7	0,0

Таблица 20

Вероятность нулевой гипотезы (p-values)

Семья	p-values		
	A2	C2	C3
A2	–	0,000000	0,000019
C2	0,000000	–	0,000000
C3	0,000019	0,000000	–

Исходя из принятого критического значения этой вероятности (уровня значимости, равного 0,05 или 0,01), можно заключить, что статистически достоверные различия по комплексу признаков, служащих компонентами продуктивности, существуют между всеми сравниваемыми группами – семьями. Близкими по комплексу признаков продуктивности оказались результаты выращивания семей А2 и С3 (расстояние Махаланобиса составило 675,2 усл. ед.), от которых значительно отличалась семья С2.

Выбор критерия отбора ценных генотипов является центральной проблемой селекции по комплексу признаков. Возможность «свёртывания» информации об объекте в значение линейной комбинации полностью не решает задачу, хотя расстояния между объектами отбора в пространстве дискриминантных функций и правомерно интерпретировать как «генетические». Действительно значения новых признаков, какими являются линейные комбинации, не могут быть переведены в масштаб селекционно значимых, и критерий отбора лучших групповых генотипов ещё следует определить.

Перспективно сравнение групп с некоторым идеальным для данных условий объектом. В определённом смысле его можно

назвать селекционной моделью, максимальное приближение к которой и составляет цель селекции. Подход можно реализовать, поместив в пространство информативных линейных комбинаций некоторую точку, координаты которой определяются с учётом заранее введённых весов селекционно значимых признаков и условий выращивания, сложившихся в ходе конкретного сравнительного эксперимента. Задача сравнения групп сводится при этом к оценке их расстояний до модели, и лучшая группа выбирается по наименьшему из них.

В самом определении «идеальный для данных условий» по существу уже сформулированы все требования к модели. «Идеальный» означает наилучший по комплексу селекционно значимых признаков, а фраза «для данных условий» требует пояснения.

Система связей признаков – категория групповая, доступная изучению только в составе группы объектов. Так, корреляционную структуру характеристик системы «объект – условия выращивания» можно изучить, лишь оперируя данными некоторой совокупности опытов, составляющих сравнительный эксперимент. Выявляемая при этом корреляционная структура показателей найдёт отражение в коэффициентах линейных комбинаций. Подстановка коэффициентов в вектор численных значений конкретного опыта наполнит этот вектор информацией о корреляционной структуре показателей, свойственной всей совокупности опытов. Здесь и открывается возможность вписать идеальный объект в условия конкретного эксперимента.

Вектор характеристик модели предлагается определять исходя из максимальных для данного сравнительного эксперимента значений признаков продуктивности, средней плотности посадки и минимальной посадочной массы рыб. Исходные данные для определения параметров модели представлены в табл. 21.

Рассмотрим реализацию данного подхода. В соответствии с обсуждаемым алгоритмом получены параметры моделей для оценки семей (табл. 22).

Процедура определения координат модели в дискриминантном пространстве основана на умножении параметров модели на значения нестандартизованных коэффициентов дискриминантных функций (Raw Coefficients) и последующего решения полученных уравнений (табл. 23).

Таблица 21

Исходные данные для построения модели

Признак	Среднее	Минимум	Максимум
pp	1338,250	915,000	1790,000
pr	410,667	304,000	759,000
mk	556,667	457,000	640,000
m0	93,000	40,000	120,000
ft	62,800	40,000	100,000

Таблица 22

Расчёт значений дискриминантных функций для модели

Признак	Коэффициенты функций		Модель	Произведения параметров на коэффициенты	
	1-й	2-й			
pp	0,00227	–0,00234	1338,25	3,041	–3,133
pr	–0,01252	–0,00434	759,00	–9,499	–3,291
mk	0,12482	–0,06243	640,00	79,885	–39,953
m0	4,95496	0,09694	40,00	198,198	3,877
ft	0,05845	–0,03007	100,00	5,845	–3,007
Constant	–531,86727	32,53723	1,00	–531,867	32,537
Координаты модели				–254,396	–12,969

Рассмотрим распределение результатов выращивания трёх семей белого толстолобика в четырёх повторностях – прудах и точки селекционной модели (рис. 30).

Анализ рис. 30 позволяет сделать заключение о том, что ближе к модели оказались результаты опытов по выращиванию семьи С2, что подтверждается множественным ранговым тестом средних значений евклидовых расстояний, которые можно получить при кластерном анализе с помощью программы STATISTICA (табл. 6).

Таблица 23

Множественный ранговый тест достоверности различий групповых евклидовых расстояний

Семьи	Средние евклидовы расстояния до модели	Ранговый тест средних		
С2	78,5	*		
А2	354,5		*	
С3	406,7			*

П р и м е ч а н и е. Расположение символа «*» на разных вертикалях отражает достоверность различий средних.

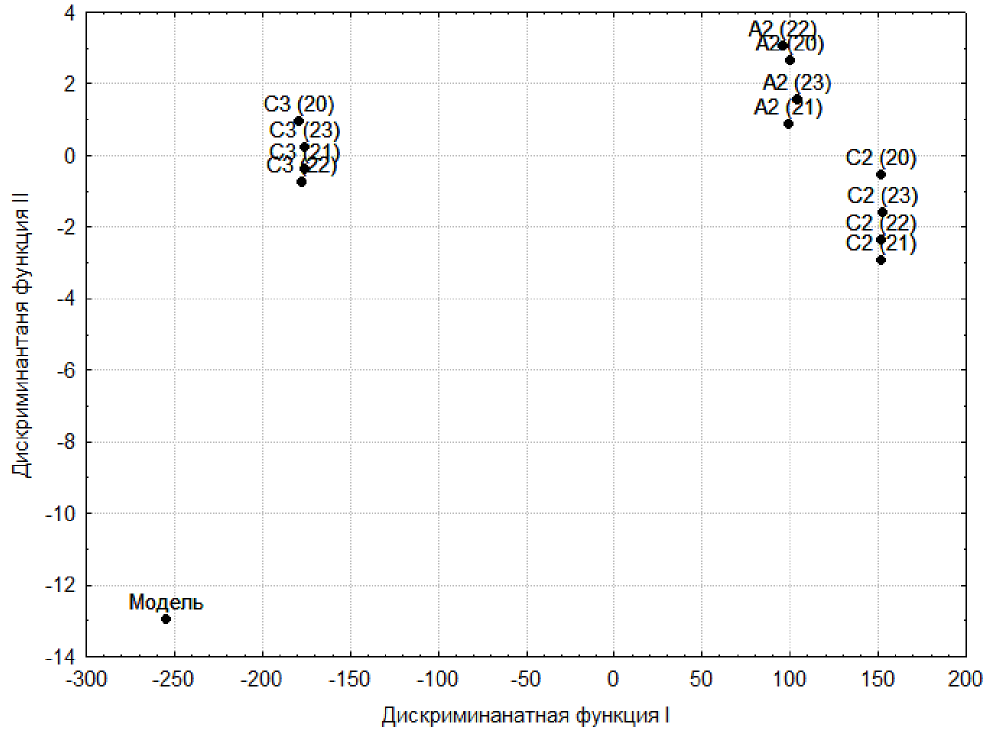


Рис. 30. Распределение точек опытов по выращиванию семей белого толстолобика и точки селекционной модели

Полученные результаты убеждают в адекватности оценки расстояния генетически различных групп (семей) до модели задаче их сравнения по продуктивности.

3.3.2. Динамика продуктивности в ряду поколений синтетической селекции местного карпа с использованием интродуцированных пород

Селекция карпа осуществлялась под руководством талантливого селекционера Александра Михайловича Бурло. Цель обсуждаемой селекционной программы заключалась в повышении показателей, являющихся слагающими продуктивности. В числе главных из них: общая рыбопродуктивность (ц/га), товарная масса при вылове (г) и выживаемость (%) (Тюрин В.В., Бурло А.М., 2009).

В качестве основной стратегии при улучшении местного карпа была принята так называемая синтетическая селекция, основанная на постепенном обогащении генетической гетерогенности исходного селекционного материала за счет гибридизации местного карпа (М) с украинским (У), а также заслужившими одобрение рыбоводов румынским (Р) и немецким (Н) карпом.

В качестве исходных компонентов скрещивания выступали чешуйчатая и зеркальная формы местного и украинского карпа. В 1981 г. были проведены первые межпородные скрещивания по схеме: местный чешуйчатый карп на украинский чешуйчатый и местный зеркальный на украинский зеркальный (Мч × Уч; Мз × Уз).

В 1985 г. было осуществлено скрещивание местно-украинских гибридов с румынским чешуйчатым и немецким зеркальным карпом (МчУч × Рч; МзУз × Нз). Цель гибридизации заключалась в повышении продуктивных свойств селекционируемого материала за счет генофонда интродуцированных пород. Третье скрещивание, реализующее принцип синтетической селекции, выполнено в 1989 г. при достижении тройными гибридами возраста четырехлеток. В качестве нового вводимого компонента вновь выступали чешуйчатая и зеркальная формы украинского карпа.

Для повышения адаптационных свойств будущей породы четвертое, завершающее синтетическую селекцию скрещивание вновь выполнено с привлечением местного карпа.

Оценка продуктивных качеств осуществлялась на втором году жизни при достижении группами рыб товарного возраста.

Каждую из сравниваемых групп выращивали в нескольких прудах, причем чешуйчатые и зеркальные формы каждой группы выращивали в условиях совместной посадки, в одном пруду. Специфика производственной деятельности рыбоводного предприятия не позволяла вычленить отдельно эффект чешуйчатой или зеркальной формы. Поэтому при анализе данных в качестве общего обозначения местного чешуйчатого и местного зеркального карпа далее будем использовать общий символ М, украинского чешуйчатого и украинского зеркального – символ У, чешуйчатой румынской и зеркальной немецкой пород – «И», т.е. импортированные породы. Общая схема селекционного процесса представлена в табл. 24.

Таблица 24

Схема синтетической селекции
путем межпородных скрещиваний местного карпа
с интродуцированными породами

Номер скрещивания	Компоненты скрещивания	Межпородный гибрид	Год постановки скрещивания
Первое	М × У	МУ	1981
Второе	МУ × И	МУИ	1985
Третье	МУИ × У	МУИУ	1989
Четвертое	МУИУ × М	МУИУМ	1993

Генетически различные группы рыб сопоставляли по комплексу из пяти показателей: средняя конечная масса рыб; выход рыбы с единицы площади пруда – рыбопродуктивность; выживаемость; средняя посадочная масса рыб; плотность посадки.

Таким образом, в общей сложности сравнению подлежало пять групп рыб с разной степенью родства: местный карп, гибрид местного и украинского карпов, тройной гибрид с участием румынского карпа, четверной гибрид, полученный при возвратном скрещивании с украинским, и конечный пятерной гибрид – от возвратного скрещивания с местным карпом.

По аналогии с многими другими биологическими исследованиями и в рыбоводстве разные пруды нередко рассматриваются как повторности опыта. Однако фактически различия условий

разных прудов и соответствующее им влияние на продуктивность столь велики, что в литературе высказывалось предложение рассматривать группу рыб в конкретном пруду не как повторность, а как отдельный эксперимент. Статистическая обработка совокупности таких экспериментов преследует цель «снятия» эффектов различных прудов с целью объективного сравнения различных групп, в нашем случае местного карпа и его гибридов различной сложности. Материалы по исследованию структуры изменчивости признаков послужили основанием для выбора именно такого подхода как наиболее корректного.

Результат анализа структуры изменчивости основных признаков-компонент продуктивности представлены в табл. 25.

Таблица 25

Результаты однофакторного дисперсионного анализа признаков продуктивности гибридов карпа с фактором «номер гибрида»

Изменчивость	Сумма квадратов	Степень свободы	Средний квадрат	Критерий Фишера	Доля в общей дисперсии, %
Средняя товарная масса тела					
Между гибридами	81762,0	4	20440,5	2,9	15,6
Остаточная	325726,0	46	7 081,0	–	84,4
Выживаемость					
Между гибридами	568,4	4	142,1	2,9	10,4
Остаточная	2988,1	46	64,9	–	89,6
Рыбопродуктивность					
Между гибридами	112,2	4	28,0	3,1	17,1
Остаточная	414,9	46	9,0	–	82,9

Для всех трёх признаков были установлены достоверные различия между местным карпом и его гибридами различной сложности: М, МУ, МУИ, МУИУ, МУИУМ. Однако следует отметить, что доля этих различий в общей изменчивости признаков была относительно невелика и варьировала от 10,4 до 17,1 %.

В изменчивости всех трех признаков очевидно доминировала остаточная дисперсия (82,9–89,6%). Ситуация вполне ожидаемая, если напомнить об известных в теории и практике рыбоводства значительных эффектах комплекса факторов, объединяемых понятием «рыбоводный режим выращивания».

Таким образом, различия факторов среды неизбежно влияют на выращиваемый материал, изменяя его характеристики. При этом продуктивность оказывается, выражаясь математически, функцией многих переменных условий выращивания. Как видно из табл. 25, их совокупный эффект столь велик, что выявление генетически обусловленных различий по продуктивности превращается в серьёзную проблему.

Для корректной оценки генетически обусловленных различий между местным карпом и его гибридами остается единственное – «снять» модифицирующие эффекты среды выращивания за счет применения специальных статистических методов.

Перспективность использования при решении последней задачи системного анализа изменчивости комплекса показателей, включающих как генетические характеристики групп рыб, так и важнейшие характеристики условий выращивания, ранее уже была доказана (Тюрин В.В., Волчков Ю.А. 1993; Тюрин В.В., Волчков Ю.А., Радецкий В.П., Сержант Л.А., 1992; Тюрин В.В., 1994; Тюрин В.В., 2008). Нашей задачей стала реализация подхода на полученном материале.

Оптимальным способом объединения признаков в линейную комбинацию является построение её методом дискриминантного анализа (Тюрин В.В., 2000). Его цель, как следует из названия, заключается в разделении групп (гибридов) на фоне искусственно минимизированной внутригрупповой дисперсии. Дисперсия между прудами и оказывается внутригрупповой, когда в дискриминантный анализ вовлекаются группы, объединяющие опыты с местным карпом или одинаковыми категориями гибридов.

В рамках дискриминантного анализа было получено искомое двумерное пространство, успешно разделяющее все пять сравниваемых групп рыб (рис. 31). Необходимо только уточнить, что каждая точка в этом пространстве соответствует отдельному варианту эксперимента, т.е. группе рыб соответствующей принадлежности в конкретном пруду.

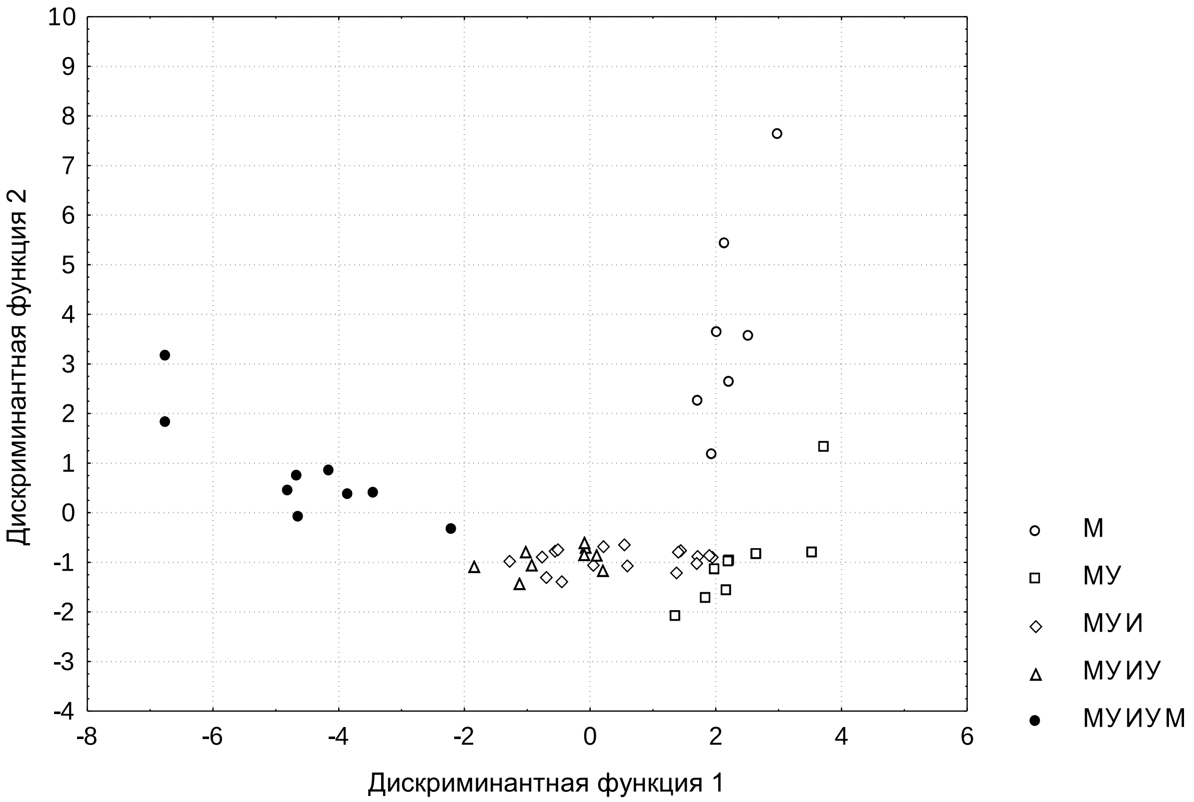


Рис. 31. Распределение опытов по выращиванию местного карпа и сложных гибридов в пространстве первой и второй дискриминантных функций

Распределение точек опытов уже прямо указывает на различия сопоставляемых групп рыб по продуктивности, но окончательный вывод может быть сделан лишь по итогам количественного анализа межгрупповых расстояний Махаланобиса, вычисляемого между центрами групп (табл. 26).

Из табл. 26 следует, что статистически достоверные межгрупповые различия по продуктивности не устанавливаются в одном единственном парном сравнении групп – тройного и четверного гибридов. Наиболее удаленными оказались местный карп и конечный продукт синтетической селекции – пятерной гибрид (61,48 усл. ед.). Это могло бы служить главным доказательством эффективности проведенной селекции, если бы не было уже доказано, что комплексно оцененная продуктивность последнего выше, чем у местного карпа и предшествующих гибридных форм.

В последнем тезисе отражена, пожалуй, как сама суть проблемы, так и эффективность ее последующего решения. Действи-

тельно, как видно из приведённых в табл. 26 расстояний Махаланобиса, дискриминантный анализ позволяет лишь эффективно разделить сопоставляемые группы, оценив расстояния между ними, но не решает проблему выбора лучшей группы.

Таблица 26

Матрица парных расстояний Махаланобиса между местным карпом и его гибридами по комплексу признаков продуктивности

	М	МУ	МУИ	МУИУ	МУИУМ
М	–	26,90	27,88	33,35	61,48
МУ	0,00	–	6,99	13,01	58,34
МУИ	0,00	0,00	–	2,00	33,16
МУИУ	0,00	0,00	0,11	–	23,11
МУИУМ	0,00	0,00	0,00	0,00	–

П р и м е ч а н и е. Выше главной диагонали матрицы приведены расстояния Махаланобиса, усл. ед.; ниже – вероятность нуль-гипотезы об отсутствии различий.

Единственно перспективным решением задачи в рамках использованного подхода является сравнение групп с некоторым идеальным для данных условий объектом – селекционной моделью, максимальное приближение к которой и составляет цель селекции.

Подход можно реализовать, разместив в пространстве дискриминантных функций некоторую точку, координаты которой соответствуют параметрам идеального объекта – селекционной модели. Параметры модели были определены в соответствии с предложенным ранее алгоритмом (Тюрин В.В., 2008).

Из рис. 32, на котором отражены центроиды групп и точки модели, следует, что наименьшее расстояние до модели обнаруживает конечный продукт синтетической селекции – пятерной гибрид, а наибольшее – исходная форма, в качестве которой выступал местный карп.

Это заключение полностью подтверждается и данными табл. 27, свидетельствующими о монотонном снижении расстояний до модели в ряду поколений синтетической селекции.

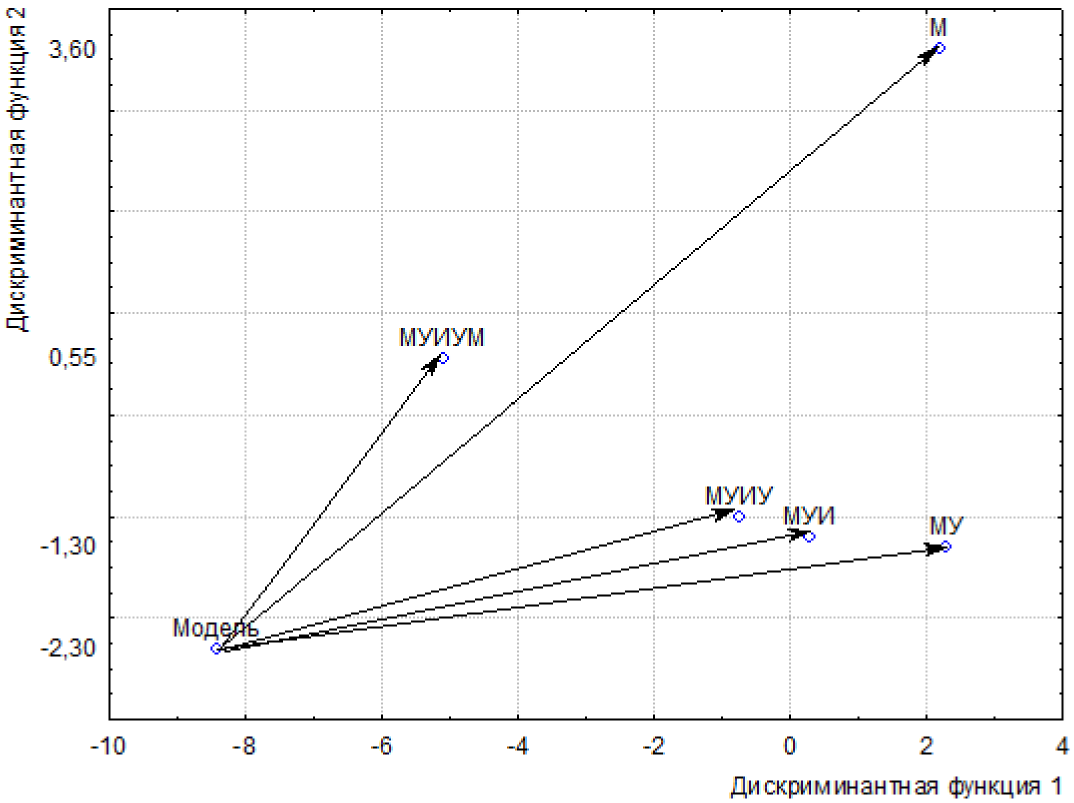


Рис. 32. Распределение центроидов групп местного карпа,
его гибридов относительно точки селекционной модели
в пространстве дискриминантных функций

Таблица 27

Итоги оценки евклидовых расстояний гибридов различной сложности до селекционной модели

Обозначение гибрида	Среднее евклидово расстояние, усл. ед.
М	12,6
МУ	11,1
МУИ	9,2
МУИУ	8,2
МУИУМ	5,4

П р и м е ч а н и е. Достоверность различия средних большинства расстояний подтверждена в дисперсионном анализе (p<0,01).

В частности, наименьшим расстоянием до цели селекции, отражённым в параметрах селекционной модели, обладает пятерной гибрид, включающий в себя генотип как исходной формы (местного карпа), так и результаты его последовательных скрещиваний с другими породами и результат возвратного скрещивания.

3.4. Использование дискриминантного анализа в биологии

Математические методы для изучения растений и животных начали активно применяться во второй половине XX в. (Любищев А.А., 1959, 1962, 1979). Так, дискриминантный анализ широко используется в различных областях биологии: ботанике, зоологии, генетике, экологии и др. В этом параграфе дан обзор научных работ по биологическим дисциплинам за последние 10 лет, которые проводились на основе дискриминантного анализа.

Н.А. Гашева (2006) изучала радикальные количественные признаки и радикальный комплекс признаков среди морфометрических параметров листьев у двух фенотипически сходных видов ив (*Salix triandra* L. и *S. pentandra* L.). Для признаков, характеризующих особенности видов, она использовала понятие «радикальные признаки» введённое Н.И. Вавиловым в 1922 г. (Вавилов Н.И., 1968) для обозначения специфических различий генетической природы. С помощью дискриминантного анализа Н.А. Гашева определяла меру сходства-различия исследуемых меж- и внутривидовых групп растений, а также фенотипические дистанции, выраженные квадратом расстояния Махаланобиса. При этом она обнаружила, что набор исследованных индексов может выступать в качестве радикального комплекса признаков для определения видов и выяснения меж- и внутривидовой фенотипической дистанции.

Вывод о том, является комплекс признаков радикальным или нет, был сделан после расчёта коэффициента трансгрессии вариационных рядов значений расстояний Махаланобиса (D^2), которые вычисляются в процессе дискриминантного анализа и служат обобщёнными показателями по комплексу использованных признаков. Модуль дискриминантного анализа компьютерной программы STATISTICA содержит два типа значений расстояния Махаланобиса (D^2): одно выражено одним числом (в случае сравнения двух групп) как показатель несходства этих исследованных групп (например, двух видов), другое – четырьмя вариационными рядами значений расстояния D^2 каждой отдельной особи сравниваемых видов от двух видовых центров (их положение определяется комплексом исследованных признаков)

как показатель несходства каждого конкретного индивидуума с групповыми характеристиками сравниваемых видов. Если вариационный ряд значений расстояний одного вида от своего центра не перекрывается с рядом значений D^2 этого же вида от центра другого вида, то комплекс признаков, примененных в дискриминантном анализе, можно считать радикальным, пригодным для различения видов. Вариационный ряд значений D^2 особей одного вида от своего же центра следует считать внутригрупповыми (внутривидовыми, внутрипопуляционными) значениями D^2, а ряд значений D^2 этого же вида от центра другого вида – межгрупповыми (межвидовыми, межпопуляционными) значениями D^2.

Результаты дискриминантного анализа, проведённого по комплексу индексов и по комплексу абсолютных значений, показали, что использованные индексы можно считать комплексом радикальных признаков, а абсолютных показателей – нет, так как в первом случае коэффициент трансгрессии внутри- и межгрупповых значений D^2 равен 0, а во втором – 70%, т.е. в случае диагностики по комплексу абсолютных показателей возможна бóльшая доля ошибки, а индексы позволяют проводить различение ив трехтычинковой и пятитычинковой по морфометрическим параметрам листьев с высокой точностью.

Значение квадрата расстояния Махаланобиса, соответствующее межвидовому уровню, значительно превышающему внутривидовой, наблюдалось и в случае исследования этим методом генеративных органов у видов, которые не родственны ивам (Гашева Н.А., 2004). При сравнении объектов из разных таксономических уровней для отнесения фенотипических различий к популяционному, видовому или надвидовому уровню имеет значение не столько абсолютное количественное выражение дистанции различий, сколько соотношение средних значений внутривидовых (внутригрупповых) и межвидовых (межгрупповых) расстояний Махаланобиса. Так, ранее проведенный нами дискриминантный анализ по двум индексам (*Cn* и *Cp*) формы семенной чешуи у двух видов ели показал на порядок меньшие межвидовые значения D^2 по своему абсолютному выражению, чем у ив, однако именно соотношение средних значений внутри- и межвидовых D^2 оказалось сходным и показывающим межвидовой уровень различий тестируемых групп.

Таким образом, Н.А. Гашева, применив методы индексов Ю.А. Филипченко (1968) и П.П. Попова (Попов П.П., 1999; Гашева Н.А., 2001) к данным морфометрических параметров листьев фенотипически близких видов ив, расчет коэффициента трансгрессии и используя дискриминантный анализ, сделала следующие выводы: 1) среди исследованных морфометрических параметров листьев ивы трёхтычинковой и пятитычинковой отсутствуют отдельные количественные радикальные признаки, однако выявлен радикальный комплекс признаков, состоящий из 26 индексов; 2) ивы трёхтычинковую и пятитычинковую можно различить, применяя дискриминантный анализ по этому радикальному комплексу признаков и статистические программы для ПК; 3) фенотипическая дистанция (D^2) между исследованными видами ив, вычисленная по комплексу индексов формы листа, оказалась равной 599; фенотипическая дистанция (D^2) при исследовании разных групп одного вида – от 2 до 8; 4) оценку степени фенотипических различий и отнесение её к меж- или внутривидовому уровню следует проводить по соотношению средних значений вариационных рядов внутри- и межгрупповых расстояний Махаланобиса (вычисляемых по результатам дискриминантного анализа), которое у сравниваемых видов ив оказалось равным 3% и практически равным по значению результатам, полученным при сравнении фенотипически сходных видов из неродственных ивам таксономических групп.

В 2008 г. Н.А. Гашева по данным морфометрического анализа листьев определяла фенотипические дистанции между основными видами ив лесной зоны Западной Сибири и сопоставляла возможности разных морфометрических комплексов для ЭВМ-диагностики исследованных видов. Ею исследованы девять видов ивы лесной зоны Тюменской области. По каждому виду сформированы смешанные выборки из летне-осенних образцов ординарных (из средней части побега) листьев от 10 и более деревьев из разных географических пунктов. Листья измерены по 10 параметрам: длина листовой пластинки (Ll); максимальная ширина листовой пластинки (Wmax); длина черешка (Lp); ширина (диаметр) черешка (Dp); расстояние от верхушки листа до линии максимальной ширины (SWmaxT); расстояние от нижнего конца листа до линии максимальной ширины (SWmaxB); ширина

листа на расстоянии 0,1Wmax от верхушки (WminT); ширина листа на расстоянии 0,1Wmax от его нижнего конца (WminB); расстояние от верхушки листа до линии ширины, соответствующей длине черешка (SLpT); расстояние от основания листа до линии ширины, соответствующей длине черешка (SLpB). Приведенные символы даны по начальным буквам соответствующих английских слов.

По 10 измерениям рассчитано 38 индексов, описывающих форму листа. Из них отдельно взято 5 индексов, которые характеризуют форму листа в соответствии с описательной методикой К. Линнея. Для описания разных частей растений К. Линней рекомендовал фактически одну и ту же схему: выделив наиболее существенную часть объекта, описать ее в системе четырех переменных: по числу, фигуре, пропорции и положению (Корона Н.В., Васильев А.Г., 2000). Форма листа описывается по этой же схеме. Индексы Линнея позволяют различать листья ив по пропорциям (Wmax/Ll) и по фигуре с учетом: а) формы контура (SWmaxT; SWmaxB); б) формы верхушки (WminT/Wmax; в) формы основания (WminB/Wmax).

Дискриминантный анализ проведен по трем комплексам (38 индексов, 5 индексов и 10 измерений) с использованием ПКП STATISTICA 6.0, в результате определены фенотипические дистанции между видами, выраженные квадратом расстояния Махаланобиса. Результаты представлены в виде классификационно-диагностической матрицы (по всем парным сочетаниям видов) и классификационно-диагностической шкалы, точкой отсчета на которой служит один вид, а по отношению к нему указываются дистанции всех остальных видов. В этом качестве взята ива черничная (*Salix myrtilloides* L.), поскольку этот вид распространен по всему евразийскому континенту в лесной зоне и прилегающих к ней территориях, он достаточно легко диагностируется как по вегетативным, так и по генеративным органам; по форме листа является серединным между крайними формами.

Классификационно-диагностические шкалы, построенные по результатам дискриминантного анализа выборок по трем комплексам морфометрических показателей листьев ив, распространенных в лесной зоне Тюменской области, имеют как сходства, так и отличия. Наиболее эффективная видовая диагностика ив до-

стигается в случае использования комплекса из 38 индексов формы листьев, так как эта модель отличается: наибольшей общей долей корректных классификаций; наибольшим показателем нижнего группового уровня корректных классификаций; наибольшими фенотипическими дистанциями между видами; наиболее дифференцированной классификационно-диагностической шкалой.

П.П. Попов (2009) изучал возможность классификации особей в промежуточной популяции ели европейской и сибирской по форме семенных чешуй с помощью дискриминантного анализа. В промежуточных популяциях ели европейской и сибирской в процессе интрогрессивной гибридизации и естественного отбора сложилось большое разнообразие особей по форме семенных чешуй, образующих непрерывный ряд. Визуальная их классификация субъективна и, следовательно, весьма условна. Дискриминантный анализ как разновидность многомерного позволяет на объективной основе разделять всю совокупность особей в популяциях на определённое число форм (Боровиков В.П., Боровиков И.П., 1997; Боровиков В.П., 1998; Иванюкович Г.А., 2010).

Благодаря этому можно получить объективную информацию о количественном соотношении фенотипов в популяции относительно «эталонных» вариантов популяций. Он определяет и величину дистанции любой особи в виде квадрата расстояния Махаланобиса (Squared Mahalanobis Distances – SMD) от «эталонных» вариантов популяций. В качестве «эталонных» или крайних вариантов популяций по величине коэффициентов сужения (*Cn*) и вытянутости (*Cp*) семенных чешуй, положению в пространстве ареала взяты выборки ели европейской из Закарпатской области Украины (г. Рахов) с территории Карпатского биосферного заповедника (370 особей) и Ивано-Франковской области (100 особей), ели сибирской – из окрестностей г. Олекминска (Саха (Якутия)) и с территории Витимского биосферного заповедника в восточной части Иркутской области. В этом случае объединённый образец включает 275 (125 + 150) особей. Для исключения случайности взяли не по одной, а по две выборки «эталонных» елей (Закарпатье, Олекминск) с близкими параметрами признака соответственно. Основными критериями подбора «эталонных» популяций были: «крайность» показателей формы семенных чешуй по соответствующим видам и достаточно большая

величина выборок. Древостои фитоценозов, где производился сбор шишек, имеют нормальный сомкнутый вид, а по лесорастительным условиям близки к лучшим в соответствующих районах.

В результате исследований П.П. Попов сделал выводы о том, что в промежуточных популяциях ели европейской и сибирской в процессе эволюции (интрогрессии и естественного отбора) образовалось большое разнообразие особей по форме семенных чешуй. Визуальная их классификация весьма условна. Дискриминантный анализ позволяет на объективной основе разделять всю совокупность особей в популяциях на то или иное число форм, определяемое задачами исследования. Полученные результаты могут быть полезны при изучении внутрипопуляционного разнообразия, в решении вопросов популяционной экологии, лесоведения.

В 2010 г. П.П. Попов изучал структуры популяций и популяционно-географической структуры сплошного ареала ели европейской и сибирской на основе биометрических показателей формы семенных чешуй (Попов П.П., 1999), дискриминантного анализа (Кулаичев А.П., 1996; Дюк В.А., 1997; Боровиков В.П., 1998), большой численности выборок и особей в них. Ель европейская и сибирская на всей изучаемой территории характеризуется относительно постепенной популяционной изменчивостью в направлении с юго-запада на северо-восток. Чёткая градация популяций по средним показателям формы семенных чешуй, изменчивость их по структуре позволяет выделить географические группы популяций с близкими фенотипическими параметрами.

Исходным материалом являются популяционные выборки шишек ели европейской и сибирской в области их сплошного ареала от Украинских Карпат, западных районов Беларуси и северо-западных районов России до Восточной Сибири. Использовался дискриминантный анализ (Дюк В.А., 1994; Боровиков В.П., 1998; Гашев С.Н., 1998) популяций по величине коэффициентов сужения (Cn – coefficient of narrowing) и вытянутости (Cp – coefficient of projection) семенных чешуй (Попов П.П., 1999). При большом числе использованных для исследования популяционных выборок (97 шт.) количество особей в каждой из них составляет, как правило, от 100 и более, а общее их число – свыше 16 тыс. шт. Поэтому все полученные результаты статистически высоко достоверны.

П.П. Попов установил, что популяции ели европейской с Карпат и ели сибирской из Восточной Сибири представляют собой виды, хорошо обособившиеся по форме семенных чешуй, поскольку между ними нет трансгрессии. Весь сплошной ареал этих видов ели чётко разделяется на пять больших групп популяций по районам. Группы I, II и IV, V разделяются ещё на два района популяций, существенно различающихся по структуре и средним показателям признака. Изучать структуру популяций ели европейской и сибирской по форме семенных чешуй следует не более чем при 5-классной градации фенотипов.

Е.Л. Зенкова (2011) с помощью кластерного и дискриминантного анализа изучала формовую структуру и географическую дифференциацию популяций ели сибирской на восточном пределе генетического влияния ели европейской по основному диагностическому признаку – форме семенных чешуй.

Ель сибирская (*Picea obovata Ledeb.*) – широко распространённый лесообразующий вид тёмнохвойной тайги в восточноевропейской части России, на Урале, в Сибири. Леса с участием ели сибирской имеют большое экологическое и хозяйственно-экономическое значение. Однако её биологические и лесоводческие особенности до сих пор изучены значительно меньше (Мамаев С.А., Попов П.П., 1989; Правдин Л.Ф., 1975), чем близкой к ней ели европейской (*Picea abies* (L.) Karst.). Район исследований включает ареал ели сибирской в западной части Тюменской области (от Салехарда до Тюмени), восточной части Свердловской области (пос. Тугулым), восточной части Республики Коми (г. Ухта, пос. Троицко-Печорск). Исходным материалом послужили 20 популяционных выборок шишек, более или менее равномерно размещённых по всей исследованной территории. От каждого дерева брали одну шишку среднего размера. Выборки включают от 100 шишек и более. Древостои по составу пород смешанные с участием главного вида от 3 до 8 единиц. Лесорастительные условия, где проводился сбор шишек, близки к лучшим в соответствующих районах. Из середины каждой шишки вырезали семенные чешуйки, на проекции которых определяли: общую высоту, ширину в наиболее широком месте (*D*), ширину на расстоянии 0,1D от верхнего края (*d*) и расстояние от верхнего края до наибольшей ширины (*h*) с точностью 0,1 мм. Показатели

D, d и *h* использовались для вычисления коэффициентов сужения (*Cn*) и вытянутости (*Cp*) верхней части чешуй (Попов П.П., Зенкова Е.Л., 2009): $Cn = d : D \cdot 100$; $Cp = h : D \cdot 100$. Эти коэффициенты позволяют объективно оценить форму семенных чешуй любой особи, ее изменчивость в популяции, а также используются в дискриминантном и кластерном анализе.

Для дискриминантного анализа в качестве эталонной популяции ели европейской взяли выборку шишек из Закарпатской области Украины (г. Рахов) с территории Карпатского биосферного заповедника (370 особей). В качестве типичной ели сибирской использовали выборки шишек из окрестностей г. Олекминска (Республика Саха) и с территории Витимского биосферного заповедника в восточной части Иркутской обл. (275 особей). Главными критериями при выборе эталонных (крайних по видам ели) популяций было предположение их видовой чистоты, т.е. отсутствие генетического влияния друг на друга и большая величина выборок.

В итоге Е.Л. Зенкова установила, что ель сибирская за Уралом, где постепенно затухает генетическое влияние ели европейской, характеризуется сравнительно небольшой изменчивостью формы семенных чешуй. При этом достаточно четко прослеживается географическая динамика признаков в направлении север-юг. Особенностью популяций вида здесь является сравнительно небольшая частота промежуточных (гибридных) фенотипов ели европейской и сибирской. Наибольшее участие их оказывается в составе популяций южных и западных частей региона. Результаты исследования могут быть полезны для решения вопросов популяционного разнообразия ели сибирской, внутривидовой систематики, выделения и использования популяционного генофонда при выращивании еловых насаждений.

Д.А. Танков (2013) рассматривал вопросы послепожарного состояния деревьев сосны обыкновенной в низкополнотных древостоях и рединах в условиях степной лесорастительной зоны Оренбургской области. Он предложил прогнозные модели послепожарного выживания отдельных особей в сосновых древостоях на основе классификационных дискриминантных функций. В качестве показателей были использованы высота дерева, диаметр дерева, высота до кроны, высота нагара.

Полученные данные по пробным площадям, относящимся к одному типу лесорастительных условий, были объединены.

Задача прогноза послепожарного состояния деревьев, которые отомрут в ближайшем будущем, очень важна для лесоводов. Прогнозирование отпада помогает определить размеры возможного пользования древесиной, изымаемой при санитарных рубках (Неповинных А.Г., 2009). Помимо оценки и прогноза отпада, не менее важно прогнозирование вероятности выживания отдельных особей в сосновых древостоях, пострадавших от лесных пожаров.

Своеобразные экологические условия предопределяют региональные особенности лесного покрова. Сильные беглые и устойчивые низовые пожары, вызывающие различные по интенсивности и продолжительности воздействия высоких температур на проводящие ткани стволов и поверхностных скелетных корней, даже при условии незначительного повреждения крон деревьев нередко приводят к гибели древостоев. В связи с этим оценка и прогноз послепожарного изменения жизненного состояния повреждённых огнём деревьев сосны обыкновенной является актуальной.

Для описания реакции дерева на воздействие пожара был применён дискриминантный анализ, использующий пространство параметров, характеризующих степень воздействия пожара на дерево и исходное состояние дерева. Все наблюдаемые объекты определяются как принадлежащие к одному из априорных классов (в данном случае – жизнеспособные и претенденты на отпад). С помощью метода наименьших квадратов определяется линейная функция от параметров максимальной высоты нагара на стволе дерева, высоты дерева, высоты до кроны и диаметра дерева, наилучшим образом разбивающая все деревья на группы в соответствии с априорными классами (т.е. жизнеспособные и претенденты на отпад). Для успешной дискриминации приходится использовать не одну, а несколько линейно независимых дискриминантных функций. Вычисление параметров дискриминантных функций было выполнено с помощью пакета статистических программ Statgraphics Centurion XVI.I. Для рассмотренной задачи достаточно использовать две функции.

Для отнесения того или иного дерева сосны в Соль-Илецком районе к определённой группе, а в нашем случае – к жизнеспо-

собным или к претендентам на отпад через некий промежуток времени были определены классификационные функции. Результаты наблюдения приписываются той группе, для которой классификационная функция имеет наибольшее значение.

В низкополнотных насаждениях сосны прогноз жизнеспособных деревьев надёжней (84,44%), чем у претендентов на отпад (86,96%). В итоге значимые показатели обеспечивают высокое качество разделения: совпадение составляет в среднем 85,6%.

Для редин, напротив, прогнозные модели для жизнеспособных деревьев сосны более надёжны (90,4%) по сравнению с претендентами на отпад (79,17%). В среднем качество разделения довольно высокое – 87,28%.

Полученные прогнозные модели показали высокую адекватность и пригодность для оценки послепожарного состояния деревьев сосны в низкополнотных насаждениях и рединах Соль-Илецкого района. Это позволит в искусственных древостоях сосны, которые подвергались низовым пожарам, более обоснованно назначать в рубку деревья, идущие в отпад. Такие меры предотвращают возможность поселения и размножения вредных насекомых, тем самым сохраняя не тронутый огнём древостой от преждевременного нападения вредителей.

И.Х. Хайров и С.В. Кабанов (2011), изучая возрастную структуру ценопопуляций осиновых насаждений юга Приволжской возвышенности, выясняли, какие экологические и ценотические факторы её определяют. С этой целью они рассматривали возрастную структуру ценопопуляций этой породы.

Признаками устойчивого положения популяции в сообществе являются полночленность и стабильность её возрастного спектра (Заугольнова Л.Б., 2000). Для изучения возрастной структуры ценопопуляций закладывались пробные площади (ПП) размером не менее 0,2 га. На каждой ПП определялись экспозиция, величина угла наклона склона, сомкнутость полога древостоя и подлеска, проективное покрытие травостоя. Учёт особей древесных видов проводился с использованием сплошного пересчёта с разделением их по видам, ступеням толщины, возрастному и жизненному состоянию. Для периодизации онтогенеза использовалась общепринятая методика (Смирнова О.В., 1976; Заугольнова Л.Б., 1988). Для того чтобы выяснить, какие экологические и

ценотические факторы определяют возрастную структуру ценопопуляций осины, был проведён пошаговый дискриминантный анализ выделенных групп онтоспектров (Пузаченко Ю.Г., 2004).

В анализ включались следующие переменные: доля участия осины, дуба и питиентных видов в составе древостоя, относительная полнота первого яруса древостоя, относительная полнота второго яруса древостоя, возраст осины, трофотоп и гигротоп условий местопроизрастания, отношение абсолютной полноты к средней высоте, высота первого и второго ярусов, сомкнутость подлеска.

Получена достаточно точная дискриминантная модель, включающая восемь переменных. Качество модели можно понять по частоте ошибочной дискриминации. Общий процент правильных отнесений к группе онтоспектров – 87%, при этом первый, третий, четвёртый и шестой кластеры распознаются на 100%. Наибольший и статистически достоверный вклад в образование групп онтоспектров вносят три переменные – сомкнутость подлеска, полнота второго яруса и трофотоп условий местопроизрастания. Такие выводы вытекают из величины частной лямбда Уилкса, *F*-критерия и вероятности нулевой гипотезы этих переменных.

На 6%-м уровне достоверности проявляется влияние доли участия в древостое дуба черешчатого. Следующие по значимости переменные – гигротоп, доля участия осины в составе древостоя, отношение абсолютной полноты к средней высоте насаждения и полнота второго яруса. На возрастную структуру ценопопуляций осины не оказывают влияния абсолютный возраст осины, высота первого и второго ярусов, доля участия патиентных видов.

Род *Epipactis* Zinn. – Дремлик – насчитывает от 60 до 80 видов, распространённых в умеренном поясе Евразии, в Африке и Северной Америке. Однако, по некоторым данным, род может насчитывать свыше 250 видов, что связано с его биологическими особенностями. В настоящее время помимо широко распространённых перекрёстно опыляющихся видов дремликов описывается большое количество самоопыляющихся автогамных и даже клейстогамных видов. Большинство из них имеет локальное распространение и, вероятно, гибридную природу.

В.В. Пинчук и В.Н. Тихомиров (2010) изучили степень изменчивости данных видов при их гибридизации и количественно оценили отличия *Epipactis atrorubens* и *E. helleborine* от гибрида *E.* × *schmalhausenii* K. Richt. Было обнаружено местообитание, в котором совместно произрастали *E. atrorubens* и *E. helleborine*. При этом наблюдалась их гибридизация. Гибриды визуально отличались от родительских видов по качественным признакам. Они имели зелёный стебель и листья (признак *E. helleborine*), но их цветки были окрашены в красный цвет, а эпихилий имел бахромчатый край и сильно морщинистые бугорки у основания (признаки *E. atrorubens*).

Дискриминантный анализ, при проведении которого в качестве группирующей переменной была взята принадлежность к определённому таксону, подтверждает выводы, полученные при сравнении средних значений изученных признаков. В ходе анализа были выделены две канонические дискриминантные функции. Первая составляет 88,8% общей дисперсии анализируемой системы, а вторая – 11,2 %. Виды и гибрид довольно чётко обособляются друг от друга в пространстве первой и второй канонических дискриминантных функций. При этом корректно классифицируется 93,5% анализируемых образцов, что является очень высоким результатом.

Если учесть признаки, имеющие достаточно высокие значения структурных коэффициентов по отношению ко второй дискриминантной функции, то эту группу признаков можно охарактеризовать как определяющую степень развитости (мощность) растений. По значению второй канонической дискриминантной функции *E. atrorubens* и *E. helleborine* не отличаются друг от друга, но отличаются от *E.* × *schmalhausenii*. Таким образом, проведённый дискриминантный анализ в целом подтверждает выводы, сделанные в ходе сравнения средних значений исследуемых признаков.

М.В. Олонова и Х.Х. Хисориев (2013) предприняли попытку выявления таксономической структуры комплекса *Poa nemoralis* L. – *Poa relaxa* Ovcz. на примере выборки, взятой в Гиссаро-Дарвазском флористическом районе Таджикистана. На основании исследования изменчивости основных дискриминаторов и применения многомерных методов статистического исследования

выборок (метод главных компонент и дискриминантный анализ) они предположили, что исследуемый комплекс представляет собой единую популяцию, состоящую из нескольких тесно связанных между собой ценопопуляций. В самой нижней, гумидной части в большинстве обитают особи, которые морфологически могут быть отнесены к *Poa nemoralis*, особи с сухого склона могут быть определены как *P. relaxa,* а значительная доля особей в переходной полосе между сухим склоном и влажной долиной – как *Poa nemoralis* × *relaxa*. При этом следует учитывать, что вся эта популяция носит гибридный характер.

Для проверки гипотезы о морфологической изоляции отдельных групп был проведён дискриминантный анализ, который используется для изучения взаимоотношений объектов с перекрывающейся у близких видов изменчивостью. Он позволяет сгруппировать множество объектов, характеризующихся набором скоррелированных входных характеристик, в несколько групп. В результате такой дискриминации в одной группе оказываются объекты, наиболее близкие по своим морфологическим признакам, тогда как межгрупповые отличия устанавливаются максимальными (Наследов А.Д., 2004; Моосмюллер Г., Ребик Н.Н., 2007; Ефимов В.М., Ковалева В.Ю., 2008). Максимальный вес при этом получают признаки с наибольшей межгрупповой дисперсией и минимальной внутригрупповой. Анализ позволил распределить весь массив данных по морфологическим группам на основании признаков V1–V23 и сопоставить результаты с предварительным распределением по выборкам 1–7. Совершенно изолированной оказалась выборка 3, высокая степень изоляции была обнаружена у выборок 4 и 5. Выборки 1, 5 и 7 перекрываются на большей части, однако заметно расхождение выборок 1 и 7. Таким образом, на основании эколого-морфологических данных исследованный гибридный комплекс *P. nemoralis* × *relaxa* в окрестностях села Джафр признать единой популяцией, состоящей из нескольких тесно связанных между собой ценопопуляций. На дне долины, по берегу ручья и в самой нижней (гумидной) части в большинстве обитают особи, которые морфологически могут быть отнесены к *P. nemoralis*, очень малая часть с сухого склона может быть определена как *P. relaxa*, и значительная доля особей, особенно в переходной полосе между сухим склоном и влажной

долиной, – как *P. nemoralis* × *relaxa*. При этом следует учитывать, что вся эта популяция носит гибридный характер.

В.М. Ефимов и др. (2009) с помощью многомерного анализа признаков у 80 образцов, принадлежащих к 14 видам амаранта, определяли принципы разделения этих видов по направлениям их использования.

Для исследования из коллекции Института цитологии и генетики СО РАН взяты 80 образцов рода Амарант, у которых анализировались следующие признаки: для зерна – форма, светлость и три оттенка окраски, масса 1000 семян, процентное содержание белка, лизина и жира; для метёлки – форма, светлость и три оттенка окраски; химические признаки: процентное содержание сухого вещества, белка, лизина, сахара, аскорбиновой кислоты в сухом веществе (листья); хозяйственные признаки: урожайность зерна, минимальный и максимальный вегетационные периоды, облиственность, минимальная и максимальная высота растений. Форма зерна кодировалась двумя значениями: 0 – округлое, 1 – округлое с бортиком. Форма метёлки также кодировалась двумя значениями: 1 – прямостоячая, 0 – все другие варианты. Цвет зерна и метёлки описывался сначала визуально, а потом каждому описанию ставились в соответствие три координаты по стандартной шкале RGB (красный, зелёный, синий). Например, цвету «коричневый» соответствуют координаты (165, 42, 42) и т. д. Сумма координат характеризует светлость (яркость) окраски, а отношение каждой цветовой координаты к их сумме – оттенок окраски. Отсутствующие значения, не более одного для каждого образца, заменялись их средними по данному признаку. Образцы с двумя и более пропусками в анализ не включались. Бинарные признаки обрабатывались вместе с количественными по формулам для количественных признаков.

На плоскости дискриминантных осей (рис. 33) видно, что зерновые формы сосредоточены в правой части рисунка, а кормовые и сорные виды – преимущественно в центре рисунка с заходом в левую часть. Таким образом, проведённый нами многомерный анализ количественных и качественных признаков в целом подтверждает неоднородность рода *Amaranthus*. Интересно отметить, что содержание сахара в листьях коррелирует с зерновым направлением, а содержание лизина в зерне – с кормовым.

Отсюда следует, что вряд ли возможен одновременный отбор на высокое содержание лизина в зерне и остальные зерновые признаки. Эта информация может быть полезна для селекционеров. Она подтверждается тем, что подрод *Amaranthus*, который наиболее интенсивно подвергается селекции, сдвинут в зерновом направлении (рис. 34). Можно предположить, что за этим стоят определённые генетические системы.

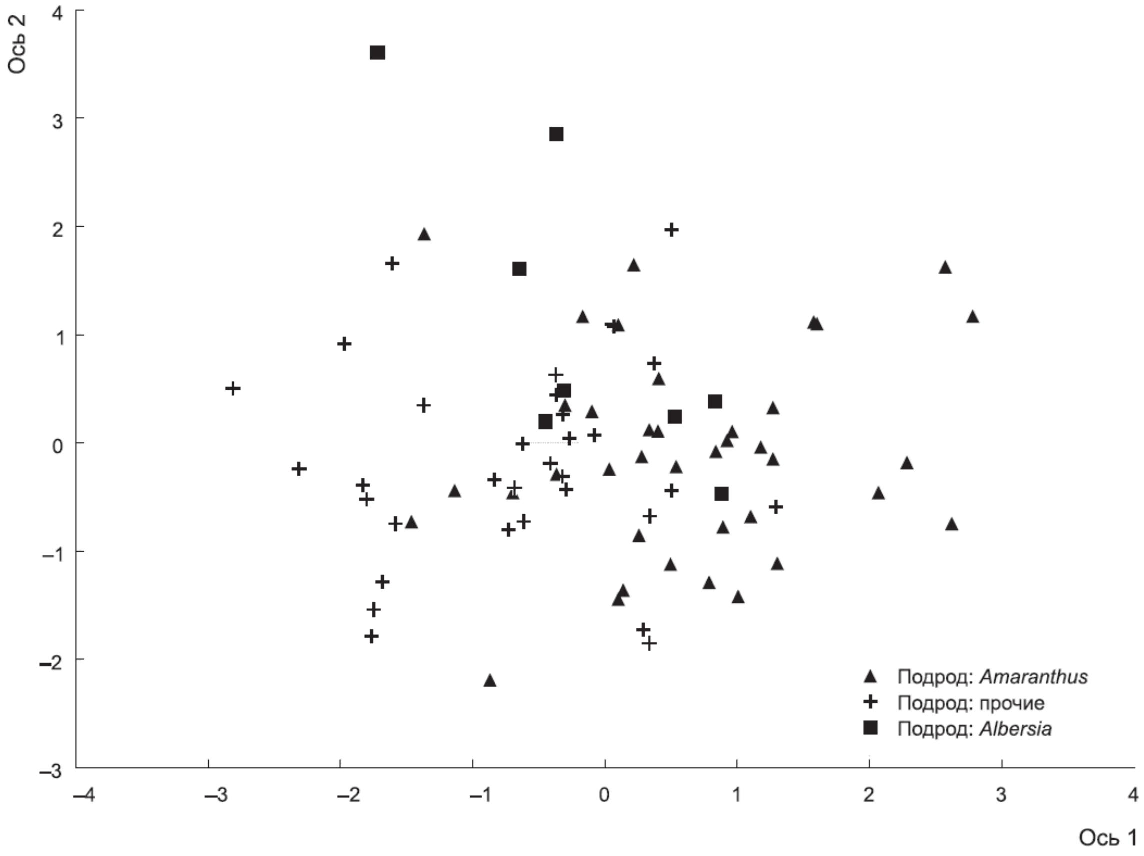


Рис. 33. Конфигурация образцов на плоскости дискриминантных осей (по Ефимову В.М. и др., 2008)

Итак, анализ рис. 33 показал, что группа зерновых амарантов характеризуется определённым набором признаков, по которым она существенно отличается от многочисленных видов и форм, условно названных кормовыми. Это можно объяснить, если мы обратимся к истории доместикации зерновых амарантов. В течение длительной селекции отбор проводился на большой габитус растений, особенно большие соцветия, которые дают огромный урожай зерна без увеличения его размера. При этом максимальное предпочтение отдавалось белым семенам, так как

белые, жёлтые и кремовые семена имеют хорошие взрывные качества и вкус. Отбирая растения с хорошими вкусовыми качествами, древние «селекционеры» неосознанно отбирали растения с высоким содержанием протеина, углеводов и масляных компонентов, что балансировало пищу с более высокой калорийной ценностью.

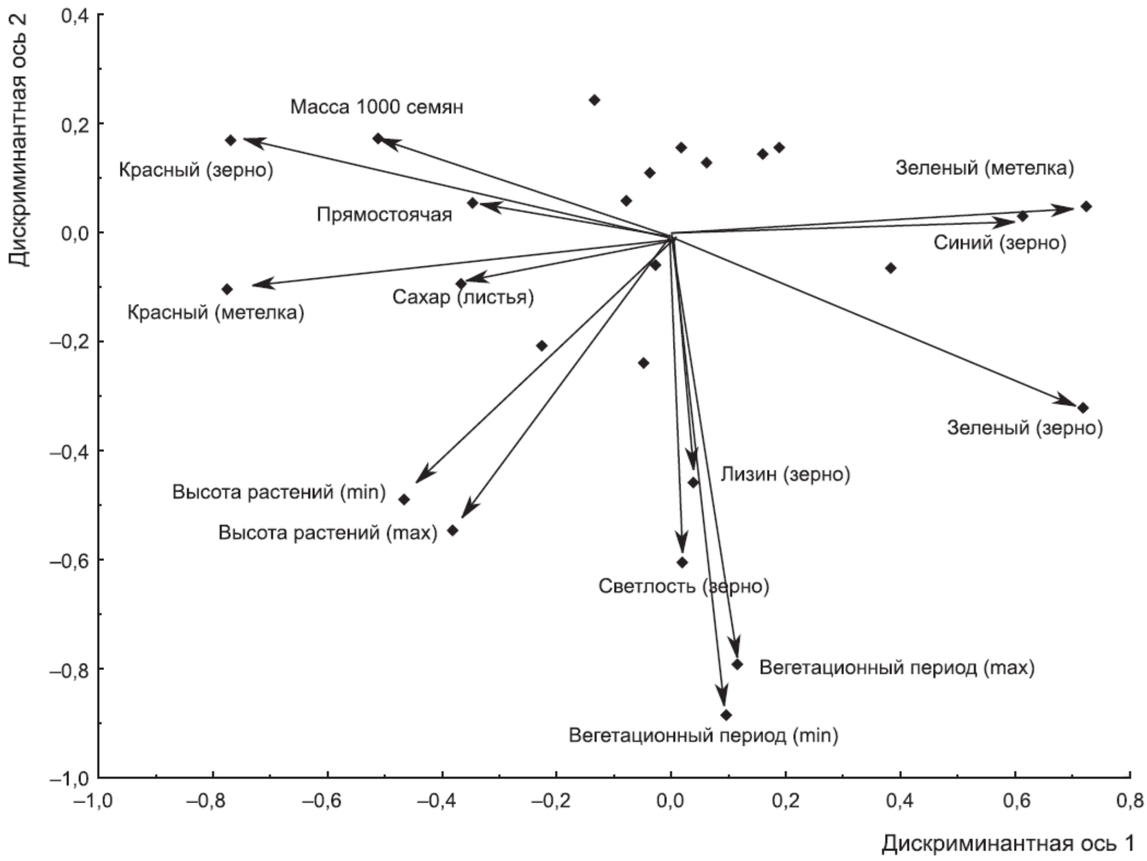


Рис. 34. Корреляции признаков с дискриминантными осями (по Ефимову В.М. и др., 2008)

Что касается красной окраски соцветий, то этому тоже есть объяснение. Древние ацтеки использовали амарант не только как продовольственную, но и как ритуальную культуру. Собираясь в завоевательные походы, воины раскрашивали свои тела краской, извлечённой из соцветий амаранта, для устрашения своего противника и привыкания к красному цвету – цвету крови. Сейчас мы знаем, что красная окраска метёлок определяется химическим соединением – амарантином.

В последние годы установлено, что амарантин является биологически активным соединением, которое воздействует на

фотосинтетические и биохимические реакции в хлоропластах и митохондриях, регулируя их на уровне клетки. Установлено также, что амарантин обладает сильными антиоксидантными свойствами, бактерицидным, антифунгицидным и общеукрепляющим действием на человека (Кононков П.Ф. и др., 2008). Не зная всего этого, древние ацтеки и инки тем не менее неосознанно отбирали растения с красноокрашенными соцветиями. Так, в результате тысячелетий сложился особый тип растений амаранта зернового направления, который можно рассмотреть с помощью многомерного анализа. На рис. 34 показано, что в первую компоненту дали вклады такие признаки, как урожайность зерна и содержание сахара в листьях, характерные для зерновых амарантов. Очевидно, что сахара, являясь продуктами фотосинтеза, депонируются в листьях, а когда наступает время цветения, они транспортируются в соцветия и используются для налива зерна, превращаясь в крахмал (зерно амаранта содержит около 60% крахмала).

Совсем другой набор признаков имеют формы амаранта кормового направления. Для них характерны высокое содержание сухого вещества, сильная облиственность, чёрная окраска зерна и зелёная окраска метёлки. Ещё один признак кормовых амарантов – тесная сопряжённость содержания лизина в листьях с облиственностью растений. По-видимому, у этих форм продукты фотосинтеза идут на образование новых листьев, в то время как у зерновых типов они в основном используются на формирование зерна. Следует подчеркнуть, что в эту группу включены в основном формы дикого типа, которых не затронуло рукотворное влияние человека.

М.О. Бурляева и М.А. Вишнякова (2010) изучали внутривидовое разнообразие чины посевной из коллекции ВИР на основе анализа фенотипических признаков. Для выявления из изученного комплекса 30 морфологических и фенологических признаков, значимых для внутривидовой классификации чины посевной, был проведён пошаговый дискриминантный анализ с последовательным исключением признаков. Использование полученной дискриминантной функции позволяет отнести отдельно взятое наблюдение (образец) к одной из этих совокупностей. Достоверные различия образцов по группам были получены по 7 призна-

кам: тип куста, окраска семенной кожуры, диаметр стебля, ширина листа и крыльев стебля, толщина боба и высота семени. Число переменных в дискриминантной функции (модели) было равно 7, значение лямбды Уилкса равнялось 0,014, т.е. мощность дискриминации была близка к 1.

При анализе расположения образцов по группам в классификационной матрице дискриминации было отмечено, что некоторые образцы, первоначально отнесенные к определенной эколого-географических группе, оказались в других группах. При рассмотрении квадратов расстояний Махалонобиса для данных случаев было обнаружено, что эти образцы находятся примерно на одинаковом расстоянии от разных групп. Например, образцы из Грузии, относящиеся к среднеевропейской группе, располагаются примерно на равных расстояниях как от средиземноморской, так и от иранской группы, более тяготея к первой. Подобная картина в распределении образцов наблюдалась и для К-800 и К-801 из Малой Азии, которые были близки и к анатолийской, и среднеевропейской группе. Такое положение данных образцов можно объяснить отсутствием значительных различий по изученным нами признакам между перечисленными группами. Возможно также, что образцы из Грузии являются гибридами образцов из иранской и средиземноморских групп.

Для выяснения закономерностей межгрупповой вариации признаков был проведён канонический анализ, в результате которого выделены две основные канонические оси. С первой канонической осью (R1) связаны признаки: тип куста, окраска семенной кожуры, диаметр стебля, ширина листа и крыльев стебля, толщина боба, а со второй (R2) – число семян в бобе и высота семени.

Таким образом, и канонический, и дискриминантный анализы показали, что значимыми признаками при классификации внутривидового разнообразия образцов являются тип куста, окраска семенной кожуры, диаметр стебля, ширина крыльев стебля, ширина листа, толщина боба и высота семени. При этом в ходе проведённого статистического анализа всех изученных признаков была выявлена их чёткая взаимосвязь с происхождением образца. По своим морфометрическим признакам образцы объединились в несколько групп (рис. 35).

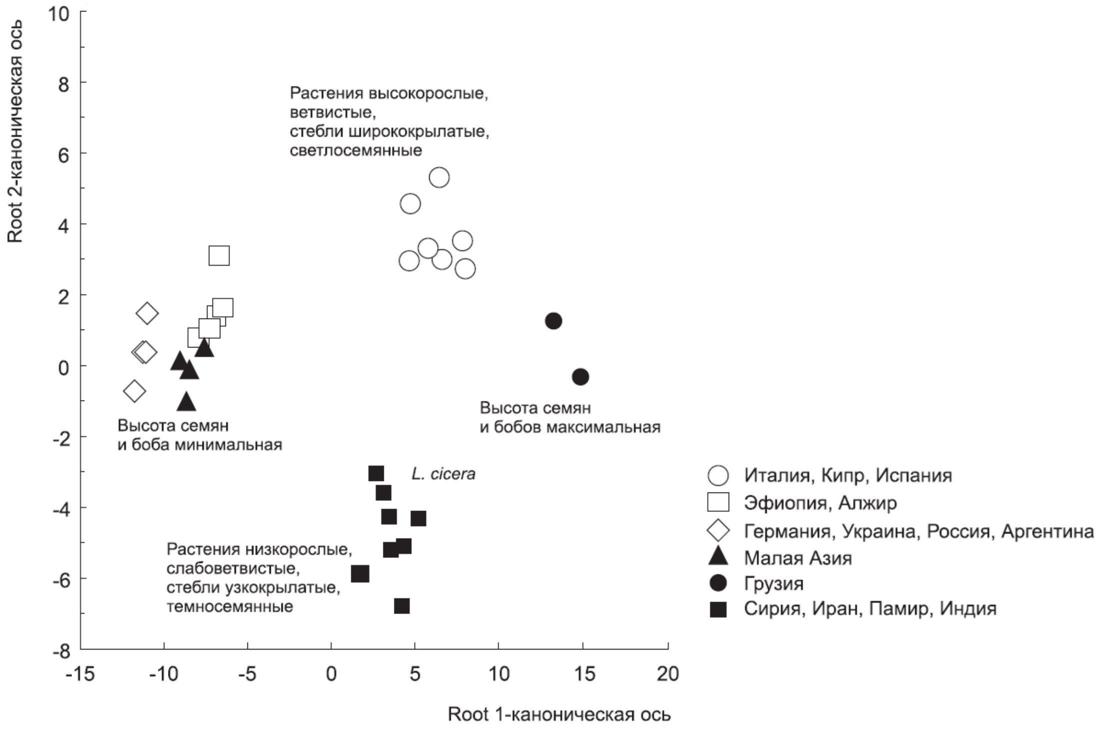


Рис. 35. Распределение образцов *L. sativus* разного географического происхождения и образца *L. cicera* в пространстве двух канонических осей по результатам морфометрической оценки (по Бурляевой М.О., Вишняковой М.А., 2010)

Образцы из Италии, Кипра, Испании и Алжира обладают максимальной высотой растений, сильной ветвистостью, самыми широкими крыльями стебля, светлыми семенами, средней высотой семян и бобов (средиземноморская, анатолийская группы). Образцы из Сирии, Ирана, Памира и Индии отличаются низкорослостью, слабой ветвистостью, узкими крыльями стебля, темными семенами и средней высотой семян и бобов (индийская, иранская группы). В левой части графика находятся местные сорта из Алжира, Эфиопии, Украины, Германии, России, Аргентины, имеющие средние показатели по параметрам: высота растения, ветвистость, ширина крыльев стебля и окраска семенной кожуры, плоские бобы и семена (анатолийская, абиссинская, среднеевропейская группы). Образцы из Грузии (среднеевропейская группа), имеющие схожий габитус растений и относящиеся к среднеевропейской группе, отличаются от последних максимальной высотой семян и толщиной бобов.

Ю.В. Чудинова (2006) оценивала показатели продуктивности сортообразцов льна в различных условиях Томской области с це-

лью выявления наиболее оптимальных условий выращивания. Исследования проводились на трёх участках: на экспериментальном участке ГНУ СибНИИСХиТ СО РАСХН (I), в дер. Половинка (II) и дер. Писаревка (III), отличающихся по комплексу микроклиматических особенностей. В качестве исходного материала она использовала 3 родительских сортообразца и 5 гибридных форм, полученных ранее по схеме согласно методике ВНИИ льна: (Т × К); (К × Т); О × (К × Т); К × (К × Т); К × (Т × К), где Т – Томский-10, О – Ольшанский, К – Крупносемянный мароканский К-6774.

Дискриминантный анализ показал, что при объединении родительских сортов и гибридов по комплексу признаков (высота растения, техническая длина, число порядков ветвления стебля, диаметр стебля, число коробочек на растении, число семян в коробочке, размеры семени) выявлены различия между участками произрастания (рис. 36).

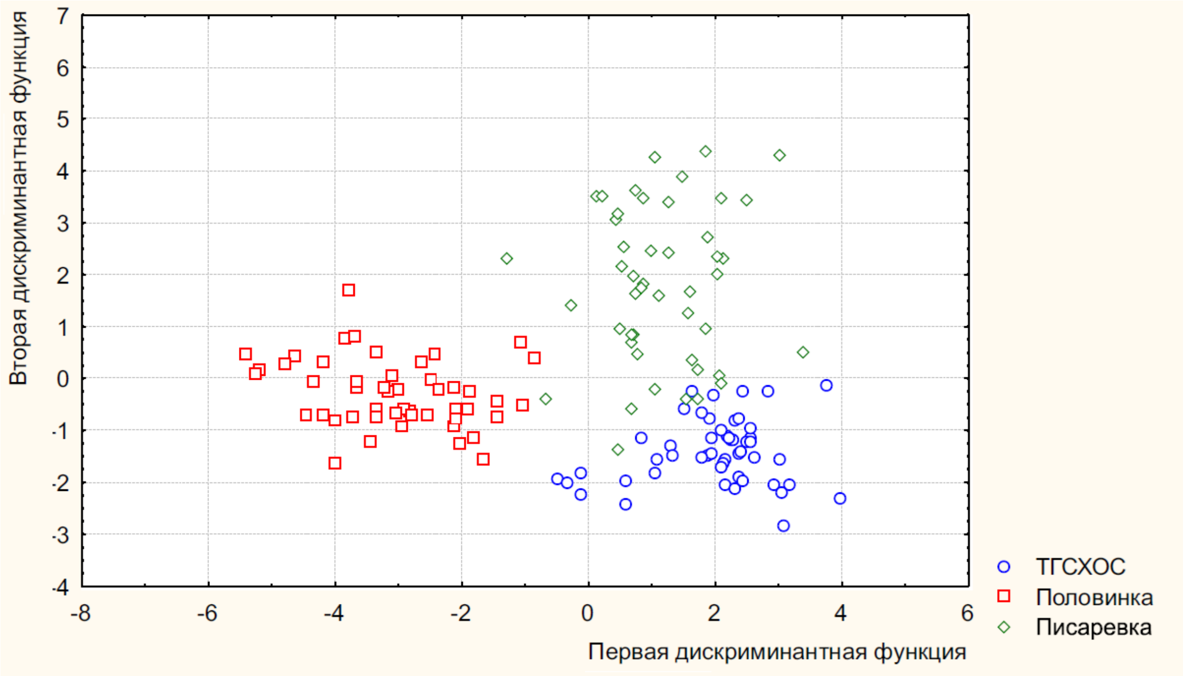


Рис. 36. График распределения признаков продуктивности
(1-я дискриминантная функция) родительских сортов Т-10, К-6774
и гибридов льна по участкам произрастания (по Чудиновой Ю.В., 2007)

Ю.В. Чудинова провела сравнительный анализ родительских сортов и гибридов по признакам продуктивности. Установлено, что в каждой области произрастания показатели продуктивности статистически достоверно отличаются между сортами и

гибридами на каждом из трёх участков. Каждый сорт и гибрид по месту произрастания также достоверно отличается от других по комплексу признаков. Она показала, что гибрид О × (К × Т) тяготеет к сорту Т и резко отличается от сорта К, родительская форма О занимает промежуточное положение между двумя образовавшимися группами (рис. 37).

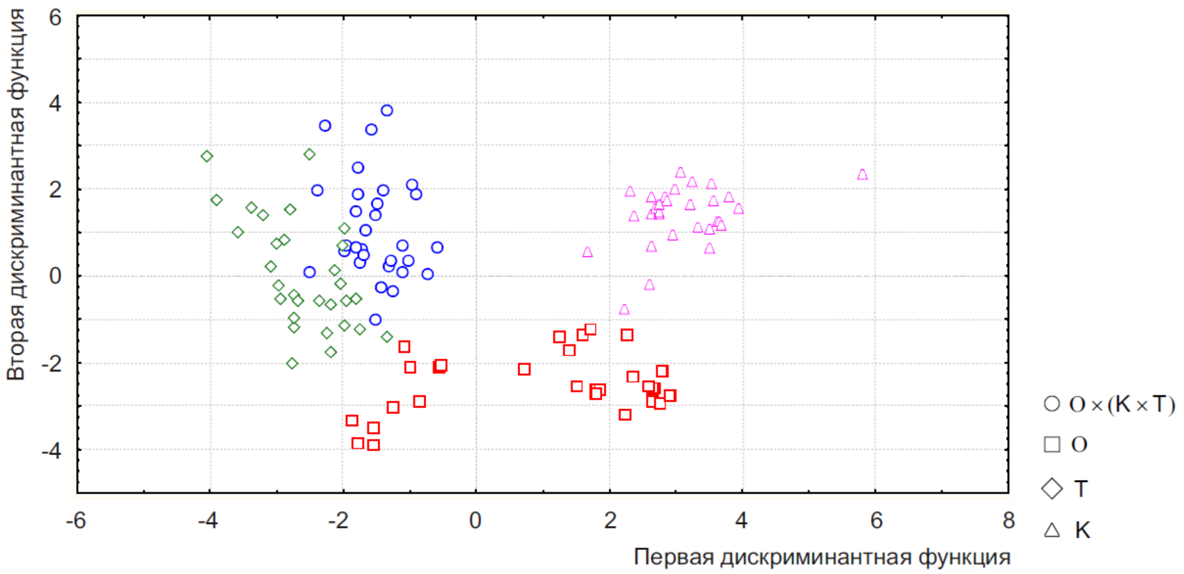


Рис. 37. График распределения признаков продуктивности (1-я дискриминантная функция) родительских сортов и гибрида О × (К × Т) (по Чудиновой Ю.В., 2007)

Гибрид К × (К × Т) занимает промежуточное положение между родительскими сортами К и Т на всех трёх участках по перечисленным признакам, но все же тяготеет к сорту К, а сами сорта резко различаются. Это говорит о преобладании генетического материала крупносемянного марроканского сорта у гибрида К × (К × Т) (рис. 38).

Дисперсионный анализ подтвердил достоверные различия между родительскими сортами и изученными гибридами по признакам продуктивности.

И.Н. Маркелов и А.А. Нижегородцев (2014) оценивали успешность классификации культурных форм примулы многоцветковой (*Primula* × *polyantha*) с помощью показателей геометрии венчика и выявляли показатели, вносящие наибольший вклад в дискриминацию этих форм.

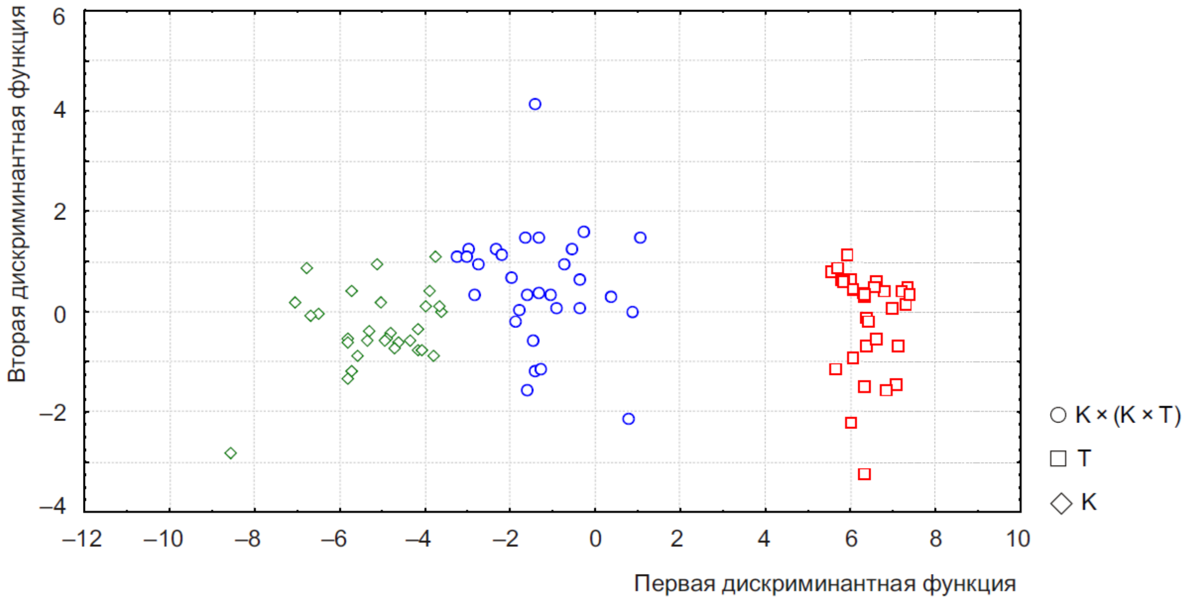


Рис. 38. График распределения признаков продуктивности
(1-я дискриминантная функция) родительских сортов
и гибрида К × (К × Т) (по Чудиновой Ю.В., 2007)

Для анализа были выбраны шесть культурных форм цветков примулы многоцветковой: *Alba*, *Atropurpuraea*, *Lutea*, *Flava*, *Rubra* и *Sanguinea*, выращенных в НИИ Ботанический сад Нижегородского государственного университета им. Н.И. Лобачевского в однотипных микроклиматических условиях. Общий объем выборки составил 324 цветка. Для оценки успешности классификации цветков *P.* × *polyantha* различных культурных форм, а также роли и вклада в дискриминацию каждого из четырёх геометрических показателей венчика был применён метод непараметрического дискриминантного анализа.

Лямбда-статистика Уилкса используется для обозначения статистической значимости мощности дискриминации в текущей модели. Границы значений лямбды Уилкса принадлежат к интервалу от 0 до 1; таким образом, значения, лежащие около 0, свидетельствуют о хорошей дискриминации, а около 1 – о плохой. В нашем случае значение лямбды составило 0,095, что говорит о хорошей разделяющей способности модели. Относительное процентное содержание собственных значений дискриминантных функций и канонические корреляции показывают, какими дискриминантными возможностями обладают функции, а также реальную полезность этих функций. Таким образом, для

графического представления данных наиболее подходят функции 1 и 2, поскольку они обладают наибольшими дискриминирующими возможностями построенной модели. С этой целью была построена диаграмма рассеяния Вороного для средних значений дискриминантных функций различных культурных форм примул (рис. 39).

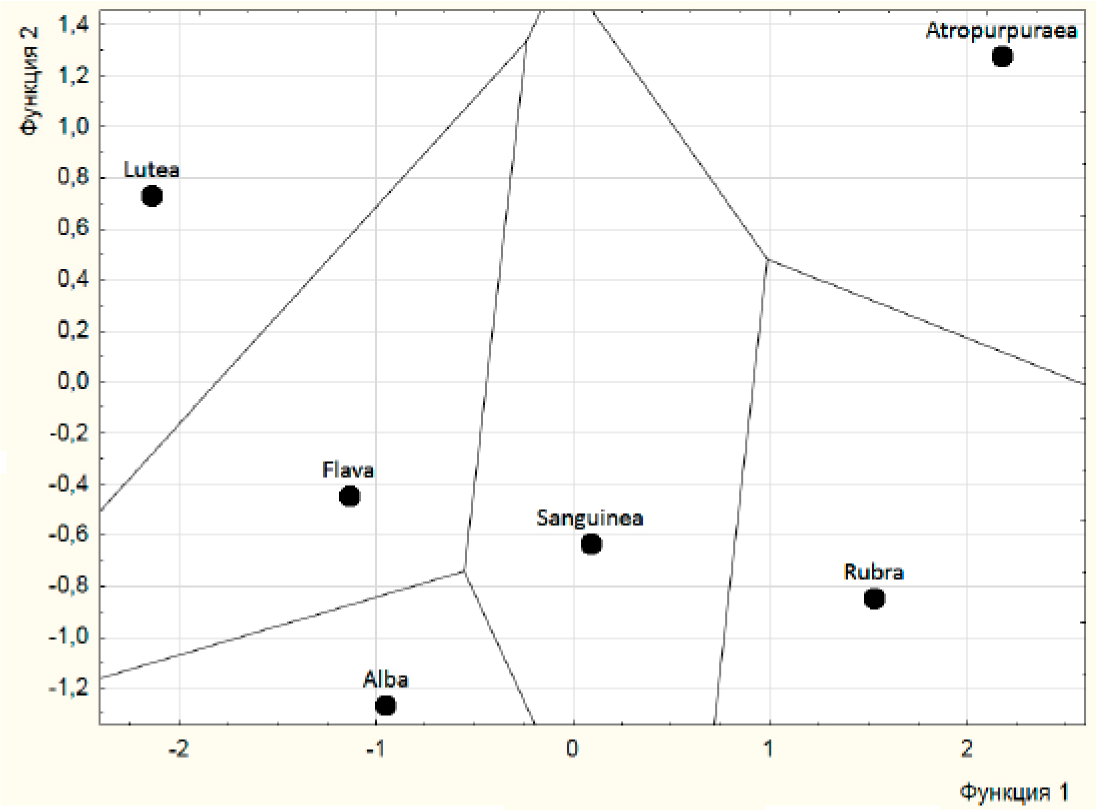


Рис. 39. Диаграмма рассеяния Вороного для средних значений дискриминантных функций различных культурных форм *Primula* × *polyantha* Mill. (по Маркелову И.Н., Нижегородцеву А.А., 2014)

Эта диаграмма позволяет определить те культурные формы, которые наилучшим образом разделяются дискриминантными функциями. Соотнося значения квадратов расстояний Махаланобиса между центроидами форм и диаграмму рассеяния (рис. 38), можно заключить, что пары форм *Atropurpuraea*/*Lutea*, *Atropurpuraea*/*Alba* и *Lutea*/*Rubra* классифицируются наиболее удачно, так как квадраты расстояний Махаланобиса между центроидами этих групп принимают наибольшие значения и в пространстве дискриминантных функций они наиболее удалены друг от друга.

Напротив, квадрат расстояния Махаланобиса между центроидами форм *Rubra* и *Sanguinea* принимает наименьшее значение, что говорит о самой слабой дискриминации среди представленных форм. Подводя итоги дискриминантного анализа геометрических характеристик венчиков цветков различных культурных форм *P.* × *polyantha*, можно заключить, что используемый набор параметров венчика цветка (ηr, ηb, Ir, Ipc) позволяет статистически значимо дискриминировать культурные формы примулы многоцветковой (лямбда Уилкса составляет 0,095). При этом наибольший вклад в дискриминацию, исходя из значений частной лямбды, вносит индекс изрезанности. Учитывая близость расположения культурных форм *Alba*, *Flava* и *Sanguinea* в пространстве дискриминантных функций, можно предположить, что они являются не отдельными культурными формами, а относятся к разным вариациям одной культурной формы. Таким образом, предложенные геометрические характеристики венчика цветка, апробированные на примуле многоцветковой, являются перспективными для задач биоиндикации.

Б.А. Абакарова и З.М. Асадулаев (2008) исследовали внутрипопуляционную изменчивость кавказского эндемика – шиповника острозубого (*Rosa oxyodon* Boiss.) – на основе комплекса количественных признаков куста и биометрических признаков листа. На территории Гунибского плато (1700 м над ур. м.) были обследованы естественные заросли (куртины) *Rosa oxyodon*. В каждой выборке рендомизированно отбирались 20 кустов. Были проанализированы следующие признаки: высота куста (см), диаметр куста (см), возраст куста, число живых побегов (шт.), число мертвых побегов (шт.), длина листа (см), ширина листа (см), длина рахиса (см), длина листочка (см), ширина листочка (см). Измеряли первый от цветка полноценный лист.

Дискриминатный анализ выявил достаточно чёткие различия между выборками по комплексу исследованных признаков. К дискриминирующим признакам относятся: возраст, длина листа, длина рахиса, ширина листочка. Суммарная точность классификации, т.е. точность отнесения к своей группе, составила 78%. Наибольшая точность (95%) отмечена для выборки №1, наименьшая (50%) – для выборки №5. Классификация позволяет разделить выборки на группы. Группа, произрастающая в более влажных местах, от-

личается крупными листьями; группа, произрастающая на открытых сухих склонах, – более мелкими размерами листьев.

Таким образом, Б.А. Абакарова и З.М. Асадулаев сделали следующие выводы: варьирование количественных признаков куста и биометрических признаков листа может быть связано с разнообразием локальных условий среды; по результатам дисперсионного анализа межгрупповые различия по всем признакам достоверны на уровне $p<0{,}05$. При этом наибольший вклад в изменчивость вносят признаки листа: длина листа $F = 45{,}3$; длина листочка $F = 39{,}3$; ширина листочка $F = 37$; признаки, имеющие большую величину индивидуальной изменчивости, менее подвержены влиянию средового фактора. Для признака «количество мёртвых побегов» CV = 103,0%, а доля влияния средового фактора – 3,80; для признака «количество живых побегов» CV = 55,3%, доля влияния средового фактора – 9,92; по данным дискриминантного анализа суммарная точность классификации составила 78%. Наиболее обособленными оказались выборки №1 (точность прогноза – 95%) и №4 (точность прогноза – 90%).

Е.Г. Бойко и др. (2011) изучали дифференциацию различных популяций ракообразных рода *Artemia* в озерах Западной Сибири на основе данных системного морфометрического анализа. Исходным материалом для морфометрических исследований послужили половозрелые самки артемии из партеногенетических популяций, выловленные в сентябре 2009 г. в озёрах юга Западной Сибири: Малое Медвежье и Большое Медвежье, Вишняковское, Невидим, Ново-Георгиевское, Большое Курейное, Требушинное и Филатово Курганской области, Эбейты Омской области, Соленый Кулат и Таузаткуль Челябинской области, Сиверга и Окуневское Тюменской области.

Помимо кластерного анализа, основная задача которого – объединение объектов в иерархическое древо, применён дискриминантный анализ. В связи с разным вкладом генетической компоненты в развитие абсолютных и относительных признаков дискриминацию провели отдельно по абсолютным признакам и индексам. Значения коэффициента лямбда Уилкса находились в интервале от 0 до 1: равнялись 0,007 по абсолютным признакам и 0,035 – по относительным. Таким образом, наблюдается чёткая дискриминация изученных популяций рачков.

Для каждой совокупности определили положение точки, представляющей средние для всех переменных в многомерном пространстве, т.е. центроиды групп. Произвели расчёт расстояния Махаланобиса от каждого центра группы, которое является мерой расстояния между двумя точками в пространстве, определяемым двумя и более коррелированными переменными. Чем меньше значение расстояния Махаланобиса, тем ближе центроиды находятся друг к другу. Анализируя полученные данные, можно констатировать, что наименьшее расстояние между центроидами рассматриваемых переменных наблюдается между популяциями рачков из озёр Б. Медвежье и М. Медвежье (1,6 – по абсолютным, 1,9 – по относительным).

Рачков из этих озёр следует считать одной популяцией вида.

Наибольшие значения расстояния Махаланобиса отмечены между рачками из озёр Б. Медвежье и Требушинное (81,5 – по абсолютным, 65,8 – по относительным). В целом значения расстояния Махаланобиса между центроидами групп по относительным признакам оказались меньшими, чем по абсолютным. На рис. 40 представлена диаграмма рассеивания центроидов по абсолютным (*а*) и относительным (*б*) признакам.

Две дискриминирующие функции вносят основной вклад в разделение изученных популяций рачков. Большее влияние оказывает функция 1 (root 1). На диаграмме просматривается четкое группирование популяций рачков из озер по солености в векторном расположении по оси *У*, поэтому, вероятнее всего, под функцией 1 подразумевается соленость среды.

Проведённый дискриминантный анализ показал, что наибольший вклад в разделение или дискриминацию изученных популяций рачков вносят такие переменные, как длина фурки, ширина абдомена, индексы ширины абдомена и длины фурки, а солёность является главной дискриминирующей функцией.

Таким образом, проведённые эксперименты позволяют сделать вывод, что дифференциация изученных популяций артемии по морфометрическим данным в большей степени основана на влиянии факторов окружающей среды, т.е. паратипической компоненте. Генетическая варианса оказывает меньшее влияние на рост рачков. Общая минерализация водоёмов значительно влияет на индивидуальные пропорции артемии. Основной дискримини-

рующей функцией является общая минерализация водоёмов. Наибольший вклад в разделение изученных популяций рачка вносят такие переменные, как длина фурки, ширина абдомена, индексы ширины абдомена и длины фурки.

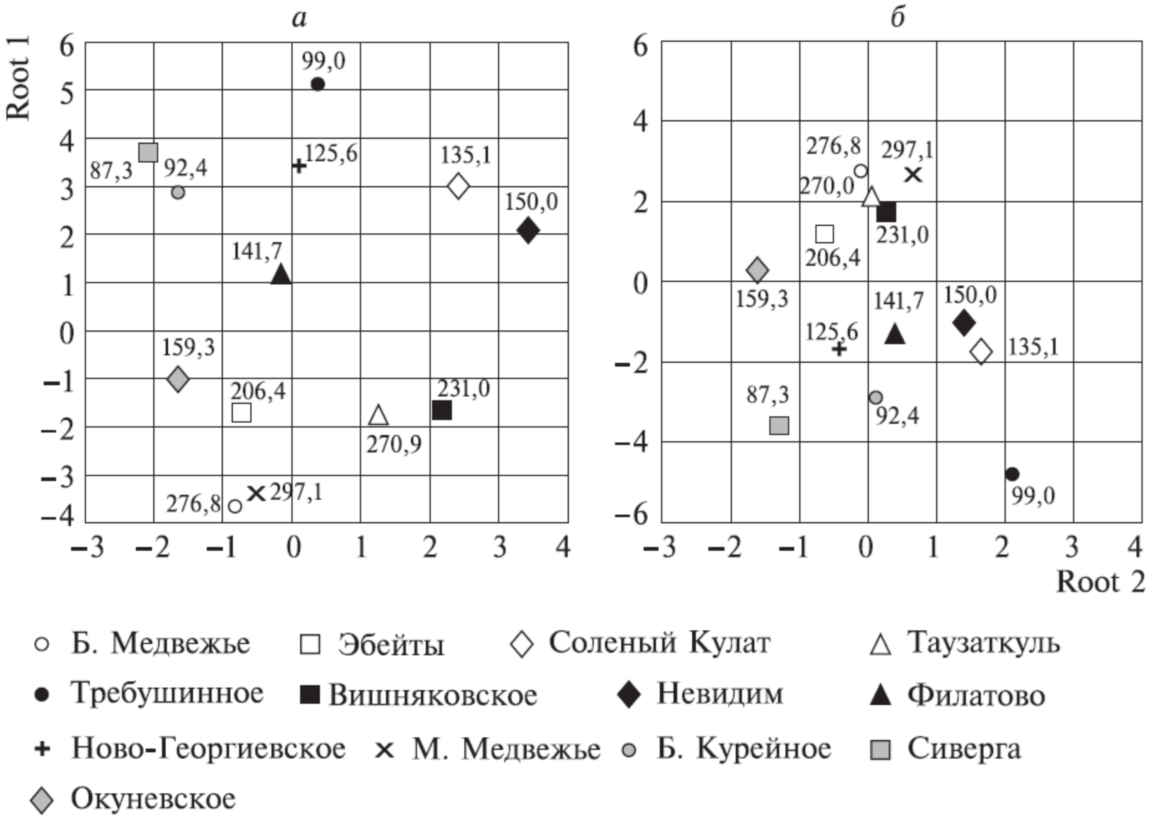


Рис. 40. Диаграмма рассеивания центроидов переменных изученных популяций артемии: *а* – по абсолютным признакам; *б* – по относительным признакам (с показателями солёности воды, г/л) (по Бойко Е.Г. и др., 2011)

М.В. Ушаков и А.И. Зиненко (2013) изучали распространение и определение подвидовой принадлежности популяций обыкновенной гадюке *Pelias berus* (Linnaeus, 1758). В 2008–2011 гг. суммарно отловлено 75 гадюк: 59 – в Воронежской области, 16 – в Липецкой. Эти исследования в регионе имеют существенное значение как с точки зрения фаунистики, так и в плане региональной природоохранной политики. Все анализы и описания внешней морфологии были проведены отдельно для самцов и самок, чтобы избежать эффекта влияния полового диморфизма, проявляющегося у обоих подвидов. При подсчёте частот встречаемости вариаций для билатеральных признаков фолидозиса изученных популяций вычислялась доля числа сторон, на которых

проявилась та или иная вариация, от общего числа сторон. То есть единицей наблюдения являлась не особь, а её половина (сторона). В этом случае объем выборки составлял 2*n*. Основанием для этого служат результаты исследований Б.Л. Астаурова (1974), установившего независимое проявление признака на разных сторонах одной и той же особи. Для отнесения выборок гадюк из Воронежской и Липецкой областей к одному из подвидов обыкновенной гадюки был проведён дискриминантный анализ.

Для классификации в ходе выполнения дискриминантного анализа используются классификационные функции. Наблюдение приписывается тому подвиду, для которого классификационная функция из пары имеет наибольшее значение. Установлено, что большинство особей отнесено к подвиду *P. b. nikolskii*. Рассчитанные средние квадраты расстояний Махаланобиса от центров выборок из Воронежской и Липецкой областей до центров «эталонных» популяций подвидов также указывают на близость экзаменуемых популяций к подвиду *P. b. nikolskii*. И это находит отражение в апостериорных вероятностях классификации изучаемых популяций как относящихся к лесостепной гадюке.

Вычисленные апостериорные вероятности классификации для гадюк из Воронежской и Липецкой областей оказались ниже, чем для «чистых» популяций гадюки Никольского. Причём наиболее низкие значения получены для самок, у которых наблюдается более сильное смещение признаков в сторону фенотипа *P. b. berus*. Проведённые исследования признаков фолидоза *P. b. berus* и *P. b. nikolskii* показали, что наибольший вклад в различия подвидов вносят число брюшных щитков и число скуловых щитков. Это позволило установить, что популяции из северной части Воронежской и Липецкой областей представлены лесостепной гадюкой, которые по сравнению с более южными «эталонными» выборками имеют большее сходство с *P. b. berus*. Сравнение ядов гадюк из Липецкой и Вологодской областей также указывает на ферментативную активность яда змей в Верхнем Подонье, свойственную для *P. b. nikolskii*. Последовательности митохондриального гена цитохрома *b*, прочитанные у особей из двух популяций в Липецкой области и одной популяции в Воронежской области, подтверждают этот вывод. Эти результаты позволяют провести границу между распространением *P. b. berus* и *P. b. nikolskii* к северу от указанных

местонахождений в Липецкой и Воронежской областей, возможно и на территории Тульской и Рязанской областей. Скорее всего, гадюка Никольского на север не выходит за пределы бассейна р. Дон, в верховьях которого и следует искать ее наиболее северные популяции. Подтверждается также приуроченность этого таксона в России к зоне широколиственных лесов в пределах Среднерусской возвышенности. Смещение изученных выборок в пространстве канонических переменных в сторону обыкновенной гадюки может свидетельствовать об интрогрессии на территории региона.

Н.Г. Борисова и др. (2013) изучали рисунок частотной модуляции звуков песен особей даурской пищухи. Различные виды млекопитающих посредством звуковых сигналов демонстрируют присутствие на занимаемом семейном или индивидуальном участке. Материалом для анализа послужили песни, записанные в четырёх колониях пищух: Склон, Ирис, Новая, Озеро.

Из сонограмм песен (не менее 20) для каждого зверька-исполнителя были сформированы выборки следующим образом: всем нотам одного типа (длительностью более 56 мс) были присвоены номера. Затем по таблице случайных чисел выбрали по десять нот для каждого зверька, сформировав обучающую выборку. Опять же с помощью таблицы случайных чисел сформировали контрольную выборку, в которую включили по десять нот для каждого индивидуально опознаваемого зверька и шесть нот из песен, записанных на тех же колониях, где авторство песен не было точно установлено (всего 46 нот). Кроме того, была создана выборка некласифицированная, куда вошли 10 нот, выбранных случайным образом из песен самца из колонии Кусты. Каждая нота была описана восемью переменными – значениями основной частоты, измеренной с интервалом 8 мс. Для выявления индивидоспецифичности выбранных звуков использовали дискриминантный анализ (статистический пакет STATISTICA 6.0 for Windows, модуль Discriminant Analysis). В качестве группирующей переменной использовали колонии.

Сначала был выполнен дискриминантный анализ данных четырёх самцов из каждой колонии. Анализ показал существование одной статистически значимой дискриминантной функции. Наибольшей нагрузкой на первую функцию обладают все переменные, кроме последней. Первые две дискриминантные функ-

ции объясняют 92,86% наблюдаемой вариабельности, из них первая – 72,68%. Хотя на плоскости двух первых дискриминантных функций точки разных групп перекрывались, центроиды групп располагались раздельно.

Затем проведена классификация нот контрольной выборки на основе двух первых дискриминантных функций, полученных при анализе обучающей выборки. 69,44% сигналов были корректно классифицированы с помощью выявленных функций.

На следующем шаге включили в анализ ноты песен самца из колонии Кусты. Корректность классификации немного уменьшилась – до 67,39%. При этом центроиды группы значений песен самца из колонии Кусты перекрылись со значениями песен самца из колонии Ирис. Это позволяет предположить, что в колониях обитал один и тот же самец, потерявший предыдущую метку.

Таким образом, установлено, что рисунок частотной модуляции звуков песен различается между особями даурской пищухи, что говорит о возможности индивидуального распознавания звуковых сигналов.

Е.В. Камалдинов (2012) определял отличия между потомками разных производителей по некоторым интерьерным показателям с последующей корректировкой полученных оценок с учетом существующих корреляционных взаимодействий.

Исследования проводились на поросятах породы СМ-1 в возрасте 19–21 дня (n = 288) и двух недель после отъёма (n = 87). Для исследований использовалась база учебного хозяйства ГПЗ «Тулинское» Новосибирского государственного аграрного университета.

Определяли массу животных, базальную температуру, частоту дыхания и сердечных сокращений, скорость оседания эритроцитов и ряд биохимических, гематологических показателей в крови сразу после ее забора: уровень эритроцитов, лейкоцитов, гемоглобина, концентрацию гемоглобина в одном эритроците, цветовой показатель, лейкоцитарную формулу, коэффициент де Ритиса, активность общей кислой фосфатазы, общей и простатической кислой фосфатазы, аланинаминотрансферазы (АЛТ), аспартатаминотрансферазы (АСТ).

Получено 5 дискриминантых функций, характеризующих поросят-сосунов (группа 1), и 3 – поросят в возрасте двух недель

после отъёма (группа 2). Коэффициенты Уилкса в целом свидетельствовали о достаточном уровне детерминации дискриминантными функциями существующих отличий потомков разных хряков по ряду интерьерных показателей. Вместе с тем следует отметить, что более высокие значения коэффициентов Уилкса установлены у поросят-сосунов. Это может объясняться наличием материнского эффекта и выражаться в изменении списка дискриминантных переменных для животных другой возрастной группы. Обнаруженные особенности обусловлены, по всей видимости, поступлением в организм животных группы 1 элементов питания с молоком матери. Для характеристики потомства хряков визуализированы канонические оценки, позволяющие судить об изменчивости комплекса установленных интерьерных показателей у поросят-сосунов (рис. 41).

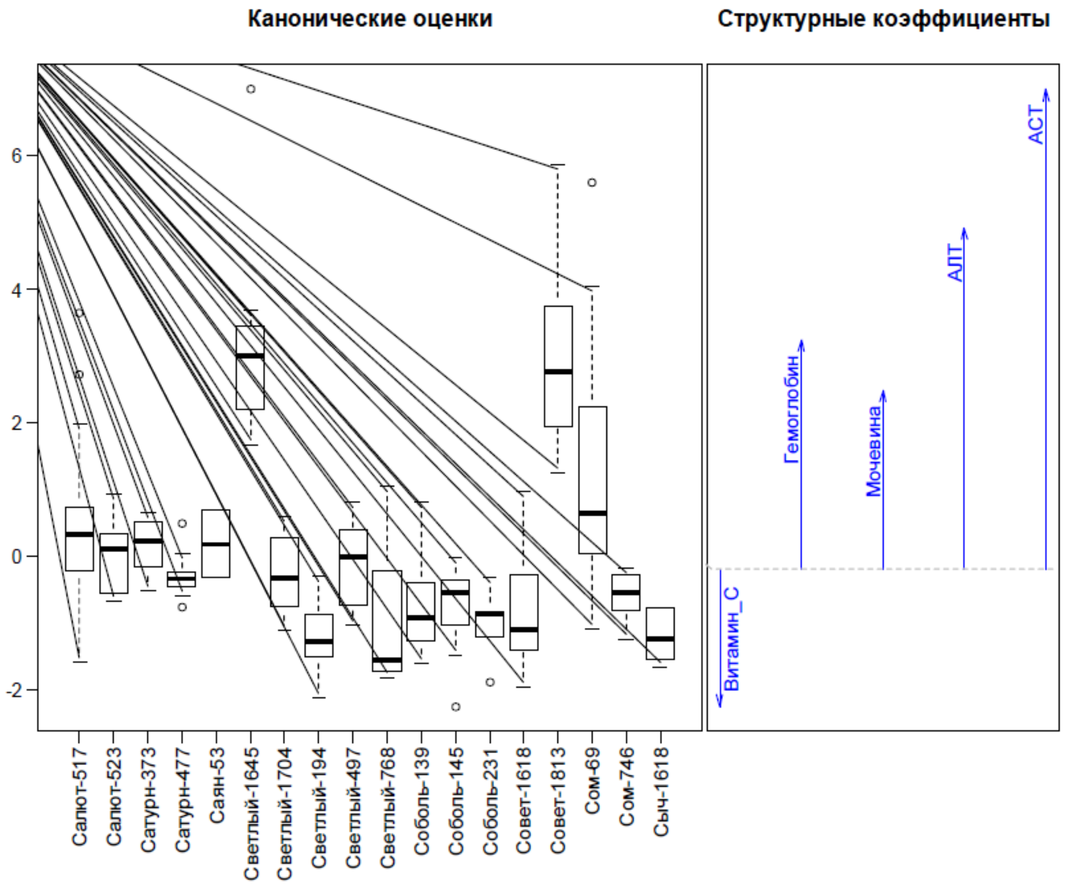


Рис. 41. Отличия потомков производителей по гематологическим и биохимическим показателям крови в возрасте 19–21 дня (по Камалдинову Е.В., 2012)

Особый интерес представляют сведения, полученные по потомкам Салюта-517, Сатурна-477, Светлого-1645, Соболя-145, Соболя-231 и Сома-69. Обнаруженные выбросы подчёркивают многомерную природу переменных дискриминантной модели и продиктованы действием как генотипа хряка, так и материнским эффектом. Наименее консолидированным оказалось потомство Салюта-517, Совета-1813 и Сома-69 по сравнению с остальными. Наряду с изложенным поросята таких производителей, как Саян-53 и Совет-1813 в наибольшей степени отличались по биохимическим параметрам от своих сверстников в подсосный период.

Проведена оценка групповых средних уровней лейкоцитов, кальция и общего белка в крови потомков некоторых хряков через две недели после отъёма. Характерной особенностью являлось отсутствие выбросов по всем изучаемым группам и изменение списка переменных канонической дискриминантной функции. С одной стороны, это объясняется нивелированием материнского эффекта группы 2 по сравнению с группой 1, где питательные элементы поступали в организм животных с молоком матери, с другой – особенностями обмена веществ растущего организма. В постотъёмный период происходит интенсивный рост мышечной и костной тканей, что объясняет наличие таких переменных, как «кальций» (Ca) и «общий белок» (ОБ), которые также сопряжены со значением дискриминантной функции. Присутствие в модели переменной «лейкоциты» (Le) объясняется ак- тивным становлением иммунной системы свиней рассматриваемой возрастной группы. Интересным представляется расположение групповых средних в верхнем и нижнем квадрантах разных периодов онтогенеза. Так, каноническая оценка потомства Саяна-53, Светлого-194, Светлого-768, Соболя-139 и Сома-746 не показала изменения знака рассматриваемых показателей. В то же время данной закономерности не наблюдалось по потомству Сыча-1618, которое также в значительной мере отличалось от средних показателей по группе. С учётом представленных данных о структурных и стандартизованных коэффициентах, предполагаемая доля влияния генотипа хряка с исключением доли, определяемой корреляционными взаимодействиями, составила по уровню лейкоцитов 0,296, об-

щего белка – 0,147. По остальным переменным подобные коэффициенты приближались к нулю. Таким образом, использование канонического дискриминантного анализа позволило выявить интерьерные признаки, по которым наблюдаются максимальные отличия с точки зрения многомерной природы их взаимодействий.

Г.П. Дробот и др. (2011) изучали гистологические особенностей мышечной ткани и лейкоцитарного состава крови лягушки озерной (*Rana ridibunda* Pallas, 1771), обитающей в биотопах с различной антропогенной нагрузкой. Объектом исследования служила мышечная ткань (скелетная и сердечная мускулатура) и кровь половозрелых особей лягушки озёрной (*Rana ridibunda* Pallas, 1771), отловленных в мае 2010 г. в трёх местообитаниях на территории Республики Марий Эл: 1) в лесопарковой зоне г. Йошкар-Олы – «Сосновая роща» (12 особей);
2) на промышленной окраине г. Йошкар-Олы (микрорайон Чихайдарово) (4 особи); 3) в Государственном природном заповеднике (ООПТ ГПЗ) «Большая Кокшага» (7 особей; контрольная группа животных).

Для изучения комплекса учтённых признаков был использован дискриминантный анализ (Дубов А.М. и др., 2000; Юнкеров В.И., Григорьев С.Г., 2002). На диаграмме рассеяния (рис. 42) видно, что по совокупности признаков отделились особи амфибий с окраины г. Йошкар-Олы. Здесь существенную роль сыграли такие признаки канонических функций, как площадь паренхимы скелетной мускулатуры (значима для первой канонической функции; $p = 0{,}032$), и доля эозинофилов (существенна для второй канонической функции; $p = 0{,}015$). Уровень их информативности составил 91%.

Таким образом, показано, что морфометрические показатели мышечной ткани и лейкоформула крови амфибий могут быть в комплексе использованы для оценки качества среды обитания. Дискриминантный анализ позволяет провести дискриминацию особей по группам исследования (местообитаниям).

Е.С. Петрачук и др. (2013) проводили анализ изменчивости морфометрических признаков леща в Обь-Иртышском бассейне

как в географически удалённых участках, так и в одной климатической зоне.

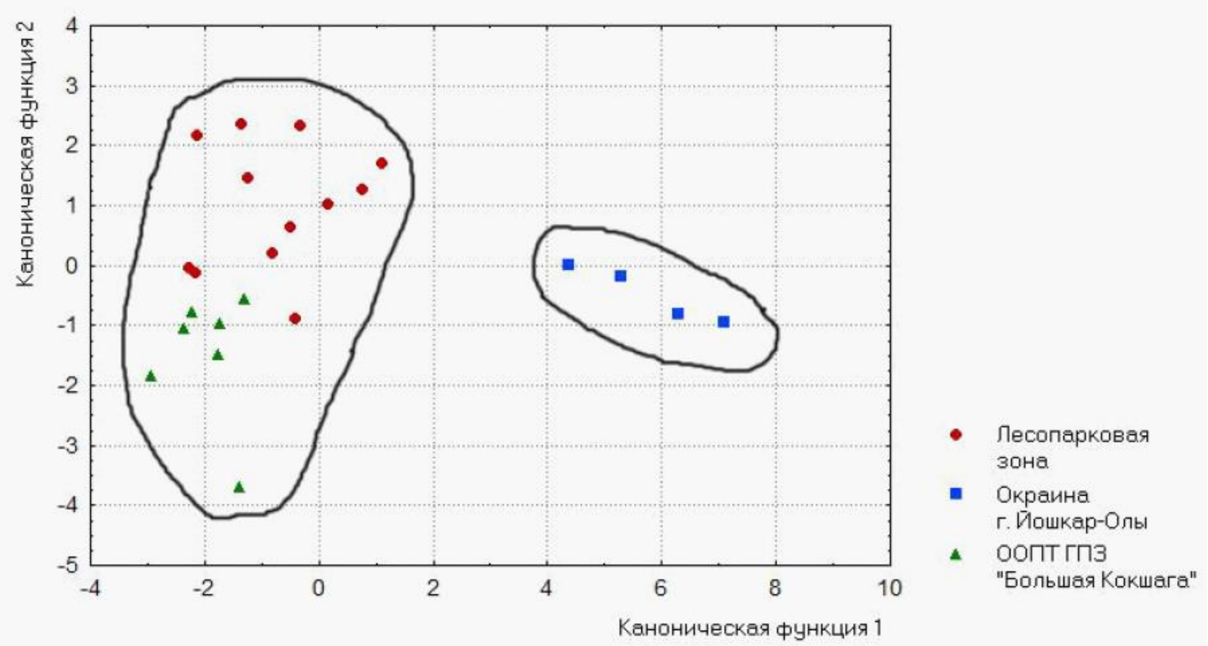


Рис. 42. Расположение особей лягушки озёрной по местообитаниям в плоскости 1-й и 2-й канонических функций (по Дробот Г.П. и др., 2011)

В проанализированных нами семи выборках леща морфологические различия самок и самцов выражены слабо, что позволило проводить дальнейший анализ без разделения по половому признаку. Известно, что меристические признаки в популяциях рыб достаточно устойчивы и генетически закреплены (Кипичноков В.С., 1987). Самые стабильные признаки леща, не имеющие варьирования, – количество колючих лучей в спинном и анальном плавниках.

Анализ изменчивости 22 пластических признаков проводили для одновозрастных особей, так как известна возрастная изменчивость пропорций тела у рыб (Мина М.В., Клевезаль Г.А., 1976). Каждая выборка характеризовалась большим своеобразием. Меньше всего достоверных различий между выборками леща из Средней Оби и оз. Монастырское (13,6 %), где наблюдались самые низкие показатели длины и массы. Влияние темпа роста на формирование пластических признаков отмечалось и другими исследователями. Более всего различается лещ из Нижней Оби и р. Пышма (77,3 %). Различия вызваны, вероятно, экологическими (кислородный и температурный режим, скорость

течения и глубина водоёмов, уровень загрязнения) и генетическими факторами.

Дискриминантный анализ позволяет диагностировать биологические объекты и явления, отличия между которыми неочевидны, с помощью уравнения дискриминации (Ивантер Э.В., Коросов А.В., 2000). В качестве обучающей выборки исследовали 161 особь леща в возрасте 5–5+ по 9 меристическим и 22 пластическим признакам. Для разделения потребовалось 18 шагов. Проведённый дискриминантный анализ семи выборок леща демонстрирует чёткое разделение на 3 морфологически обособленные группы (рис. 43).

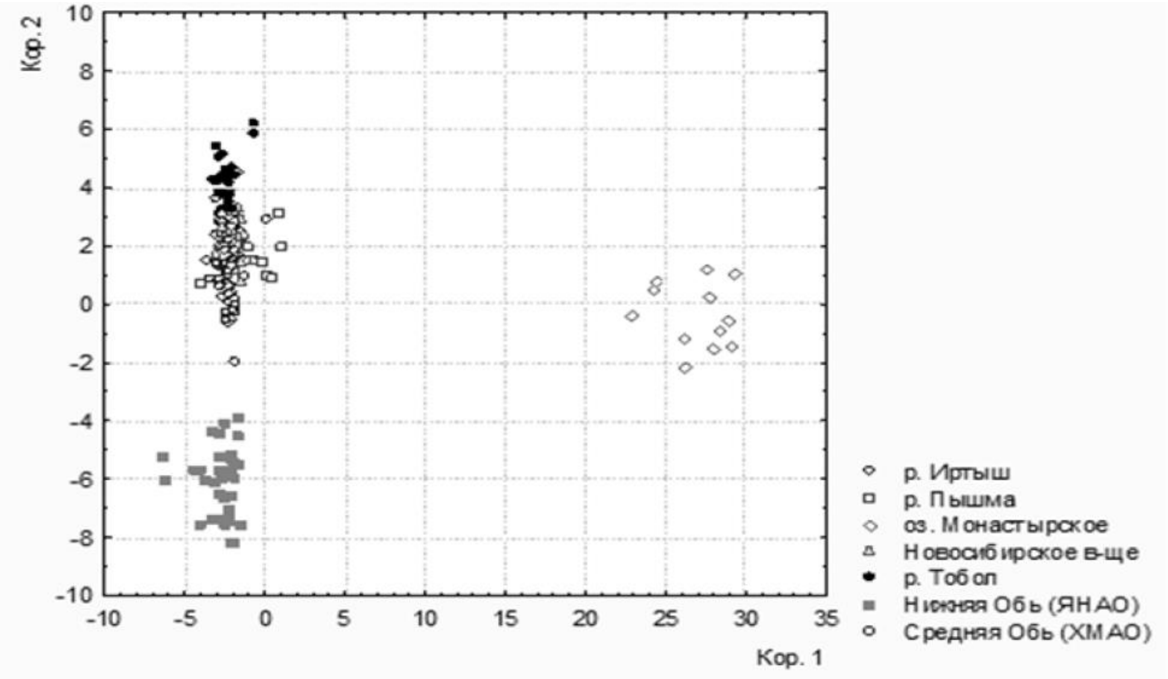


Рис. 43. Дискриминантный (канонический) анализ морфометрических признаков леща в возрасте 5–5+ различных водных объектов Обь-Иртышского бассейна (по Петрачук Е.С. и др., 2013)

При классификации обучающей выборки было 3 ошибки (2%). Статистический показатель анализа лямбда Уилкса был менее 0,001, что свидетельствуют о хорошей дискриминации групп на основе имеющихся признаков (Боровиков В.П., 1998). В тестовой выборке возраста 4+6+ из 178 особей доля переопределенных особей была велика – 14,0%. Возможно, сказывается влияние возрастных изменений, так как при определении рыб возраста 4–4+ отмечено максимальное число ошибок (23 особи из 114, или

20,2%), тогда как в тестированной выборке особей возраста 6–6+ число ошибок составило всего 3,1% и только для особей р. Тобол. Дискриминантный анализ леща из географически удалённых участков Обь-Иртышского бассейна выявил изолированность ямальской популяции. Достоверны дискриминантные функции по 10 признакам: общее число чешуй в боковой линии и прободенных чешуй в боковой линии, число чешуй над и под боковой линией, диаметр глаза, ширина лба, высота головы у затылка, высота спинного и анального плавников, ширина тела.

Дискриминантный анализ совокупности морфометрических признаков леща в одной климатогеографической зоне Обь-Иртышского бассейна выявил, что речные популяции образуют одну, но довольно подразделённую группу и полностью изолированы оказываются особи оз. Монастырское. Достоверны функции дискриминации 6 признаков: число лучей в грудных и брюшных плавниках, максимальная высота тела, длина и высота анального плавника и антепектральное расстояние. Эти признаки, за исключением высоты анального плавника, практически не значимы при дискриминантном анализе географически удалённых выборок но что вполне логично, почти все, кроме антепектрального расстояния, значимы при дискриминации всех семи проанализированных группировок леща.

Е.В. Чащиным и др. (2004) на основе анализа изменчивости метрических признаков нижней челюсти степного сурка проведена проверка предположения об адекватности результатов, получаемых при использовании не дифференцированных по полу выборок.

Анализировали две выборки байбака (*Marmota bobak* Mull.) из окрестностей н.п. Ерментау (далее – Ерментау) и н.п. Богембай (далее – Богембай) Акмолинской области Казахстана и из окрестностей н.п. Бреды (далее – Бреды) Челябинской области Российской Федерации.

Необходимо отметить, что распределение переменных во всех используемых для анализа группах приближается к нормальному. В наибольшей степени сказанное относится к не дифференцированным по полу группам вследствие их больших размеров.

Указанными различиями были предопределены результаты дискриминантного анализа проведённого для десяти выделенных групп, по 10 метрическим признакам нижней челюсти (рис. 44).

На рис. 44 видно, что центроиды исследуемых групп животных из разных частей ареала в пространстве двух первых канонических осей, описывающих 84,9% изменчивости, образовали три хорошо различимых «облака». Первое – 3 группы серых сурков из Алтайского края. Второе – 3 группы байбаков из Челябинской области РФ. Третье – 4 группы байбаков из Акмолинской области Казахстана. Положение центроидов групп байбаков из Богембая и Ерментау, образующих единое «облако», свидетельствует о принадлежности этих групп к одной популяции. Напротив, положение центроидов групп байбаков из Казахстана и Челябинской области говорит о значительных различиях, имеющих межпопуляционный характер. Заметная удалённость центроидов групп серых сурков и байбаков вдоль первой канонической оси подтверждает видовой характер отличий между ними. На это указывает также противоположный характер расположения центроидов групп самцов и самок у серых сурков и байбаков вдоль второй канонической оси. Причём у групп серых сурков половые отличия по первой оси выражены очень слабо и, напротив, хорошо заметны по второй оси. В то же время у сурков из Бредов центроиды групп самцов и самок относительно первой оси заметно расходятся, а вдоль второй оси – сближены.

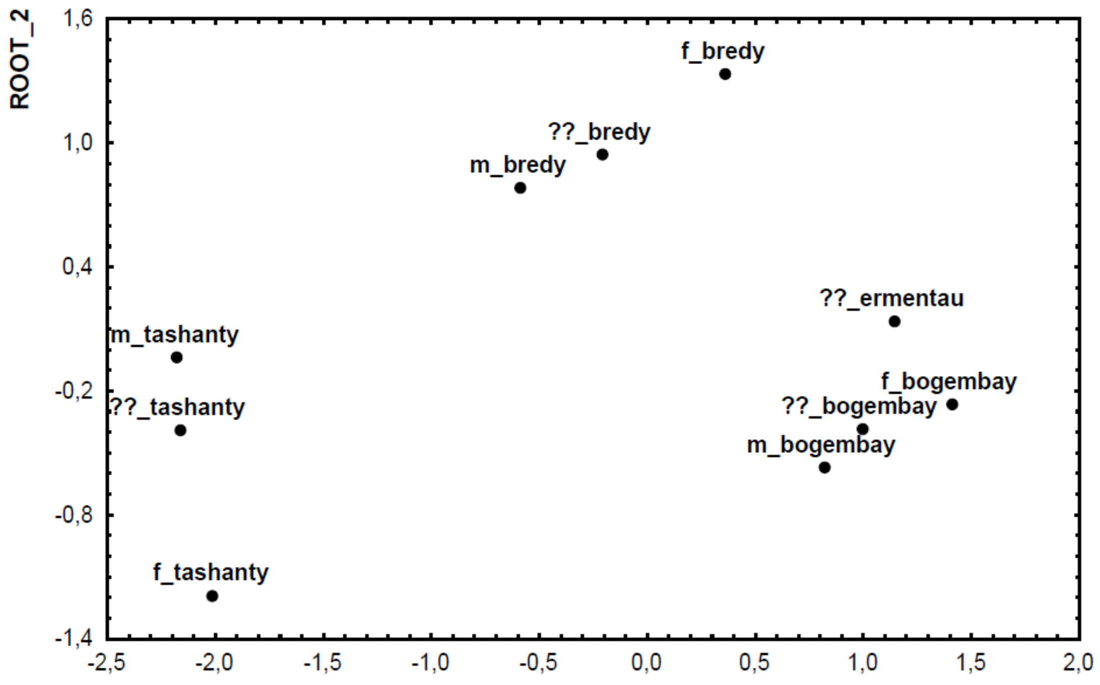


Рис. 44. Положение центроидов 10 групп сурков в пространстве двух первых канонических осей (ROOT 1) и (ROOT 2) (по Чащину П.В. и др., 2004)

Вдоль первой канонической оси центроиды групп байбаков из Челябинской области занимают промежуточное положение относительно центроидов групп серых сурков (Алтай) и байбаков из Северного Казахстана (Акмолинская область), хотя на самом деле расстояние между местонахождениями Ташанты и Бреды в два раза больше, чем между местонахождениями Ташанты и Богембай. Такой характер расположения центроидов отражает размерные особенности байбаков из Южного Зауралья.

Во всех трёх дифференцированных по полу выборках центроиды смешанных по полу групп сурков занимают промежуточное положение относительно центроидов групп самцов и самок. При этом различия между не дифференцированными по полу группами сурков такие же чёткие, как и между группами самцов или самок, относящихся к разным видам или популяциям. Такой результат доказывает полную адекватность результатов, получаемых на материале как дифференцированном, так и не дифференцированном по полу.

Центроиды смешанных групп ташантинских, богембайских и брединских сурков в пространстве двух первых канонических осей тяготеют к центроидам соответствующих групп самцов (рис. 44). Это отражает имеющееся соотношение самцов и самок в выборках, приблизительно равное 2 : 1. Надо отметить, что соотношение полов в выборках скорее всего не отражает реальную половую структуру популяций. Она вызвана, вероятно, избирательностью промысла (вследствие меньшей осторожности самцов).

Заключение

Подводя итог проведённого обзора работ можно сделать вывод о широком спектре биологических задач, которые способен решать дискриминантный анализ.

Показана возможность дискриминантного анализа выявлять объективные межгрупповые (например межпопуляционные) различия за счёт искусственной минимизации внутригрупповой изменчивости. В случаях доказательства различий можно оценивать их меру, используя расстояние Махаланобиса.

Дискриминантный анализ позволяет сократить исходный комплекс признаков за счёт исключения малоинформативных переменных. При этом открывается возможность уменьшения трудовых затрат на измерения и снятия эффекта статистического шума, мешающего сравнению групп. В дискриминантном анализе для этой цели следует использовать пошаговые процедуры, либо сравнение стандартизованных коэффициентов.

Отдельная проблема биологических исследований – определение групповой принадлежности неизвестных объектов. Это особенно важно в селекционно-генетических исследованиях, где актуальна задача идентификации селекционных достижений. Использование классифицирующих функций позволяет проводить процедуры классификации с обучением по результатам исследования обучающей выборки.

При помощи специальной методики можно получить биологическую интерпретацию межгрупповых различий, выявляемых в рамках дискриминантного анализа.

Подход основан на определении параметров идеального объекта – модели, вычислении его координат в пространстве линейных комбинаций и расчете расстояний до него от центроидов групп. Параметры модели должны соответствовать цели эксперимента, максимальное приближение к которой и должно являться критерием отбора. Именно внедрение модели в пространство, разделяющее группы, привносит биологический смысл в безразмерные линейные комбинации.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Абакарова Б.А., Асадулаев З.М.* Внутрипопуляционная изменчивость признаков *Rosa oxyodon* Boiss. на Гунибском // Юг России: экология, развитие. 2008. №4. С. 42–46.

2. *Авдеев В.Н., Ермаков Б.С., Саустов В.В.* Биологические особенности облепихи // Облепиха. М.: Лесная промышленность, 1985. С. 12–35.

3. *Айвазян С.А. [и др.]*. Прикладная статистика: Классификация и снижение размерности: Справ. изд. / под ред. *С.А. Айвазяна*. М.: Финансы и статистика, 1989. 607 с.

4. *Алексахин С.В. [и др.]*. Прикладной статистический анализ: учеб. пособие для вузов. М.: ПРИОР, 2001. 224 с.

5. *Аренс Х., Лейтер Ю.* Многомерный дисперсионный анализ. М.: Финансы и статистика, 1985. 230 с.

6. *Астауров Б.Л.* Исследование наследственного изменения галтеров у *Drosophila melanogaster* Schin. // Наследственность и развитие. М.: Наука, 1974. С. 9–53.

7. *Афифи А., Эйзен С.* Статистический анализ: Подход с использованием ЭВМ. М.: Мир, 1982. 488 с.

8. *Бербанк Л.* Избранные сочинения. М.: Изд. иностр. лит., 1955. 713 с.

9. *Берг Р.Л.* Корреляционные плеяды и стабилизирующий отбор // Применение математических методов в биологии. Л., 1964. Вып. 3. С. 24–60.

10. *Бойко Е.Г. [и др.]*. Популяционная дифференциация жаброногих рачков рода *Artemia* Leach., 1819 // Сибирский вестник сельскохозяйственной науки. 2011. №2. С. 68–75.

11. *Борисова Н.Г., Старков А.И., Руднева Л.В.* Индивидоспецифичность звуковых сигналов у даурской пищухи // Вестник Бурятского государственного университета. 2013. №4. С. 94–96.

12. *Боровиков В.П.* Популярное введение в программу STATISTICA. М.: Компьютер-Пресс, 1998. 267 с.

13. *Боровиков В.П., Боровиков И.П.* STATISTICA – Статистический анализ и обработка данных в среде Windows. М.: Филинъ, 1997. 608 с.

14. *Бурляева М.О., Вишнякова М.А.* Фенотипическое и генотипическое разнообразие *Lathyrus sativus* L. из коллекции ВИР //

Вавиловский журнал генетики и селекции. 2010. Т.14, №4. С. 769–782.

15. *Вавилов Н.И.* Закон гомологических рядов в наследственной изменчивости // Классики советской генетики (1920–1940). Л.: Наука, 1968. С. 9–50.

16. *Волчков Ю.А.* Системный анализ изменчивости в селекции рыб: автореф. дис. … д–ра биол. наук. СПб., 1994. 50 с.

17. *Гашев С.Н.* Статистический анализ для биологов: Пакет программ «STATAN». Тюмень: Изд-во ТюмГУ, 1998. 51 с.

18. *Гашева Н.А.* Межвидовые фенотипические дистанции ив по комплексам морфометрических признаков листа // Вестник экологии, лесоведения и ландшафтоведения. 2008. №8. С. 116–123.

19. *Гашева Н.А.* Опыт применения дискриминантного анализа для различения фенотипически сходных видов ив // Вестник экологии, лесоведения и ландшафтоведения. 2006. №6. С. 123–130.

20. *Гашева Н.А.* Опыт применения кластерного и дискриминантного анализа при описании структуры популяции ели по форме семенной чешуи // Вестник экологии, лесоведения и ландшафтоведения. 2001. Вып. 2. С. 113–116.

21. *Гашева Н.А.* Структура популяции ели сибирской, определяемая по радикальным признакам в разных эколого-географических условиях Северного Урала: дис. … канд. биол. наук. Тюмень, 2004. 163 с.

22. *Дерябин В.Е.* Многомерная биометрия для антропологов. М.: Изд-во МГУ, 1983. 226 с.

23. *Дробот Г.П., Мальцева Н.Л., Ведерникова А.А.* Ответная реакция некоторых тканей лягушки озерной (*Rana ridibunda* Pallas, 1771) на антропогенную нагрузку // Вестник Оренбургского государственного университета. 2011. №12 (131). С. 65–67.

24. *Дубров А.М., Мхитарян В.С., Трошин Л.И.* Многомерные статистические методы: учебник. М.: Финансы и статистика, 2000. 352 с.

25. *Дюк В.А.* Компьютерная психодиагностика. СПб.: Братство, 1994. 364 с.

26. *Дюк В.А.* Обработка данных на ПК в примерах. СПб.: Питер, 1997. 240 с.

27. *Ермолаев-Томин О.Ю.* Математические методы в психологии: учебник. М.: Юрайт, 2014. 511 с.

28. *Ефимов В.М., Железнова Н.Б., Железнов А.В.* Многомерный анализ изменчивости морфологических, химических и хозяйственных признаков в роде Амарант (*Amaranthus* L.) // Вавиловский журнал генетики и селекции. 2009. Т.13, №4. С. 811–818.

29. *Ефимов В.М., Ковалева В.Ю.* Многомерный анализ биологических данных: учеб. пособие. Горно-Алтайск: Горно-Алтайский гос. ун-т, 2007. 75 с.

30. *Животовский Л.А.* Проблема анализа комплекса признаков // Экологическая генетика и селекция. Кишинёв, 1987. С. 117–134.

31. *Заренков Н.А.* Лекции по теории систематики. М.: Изд-во МГУ, 1976. 140 с.

32. *Заугольнова Л.Б.* Оценка и сохранение биоразнообразия лесного покрова в заповедниках Европейской России. М.: Научный мир, 2000. 196 с.

33. *Заугольнова Л.Б.* Ценопопуляции растений (очерки популяционной биологии). М.: Наука, 1988. 184 с.

34. *Зенкова Е.Л.* Популяция ели сибирской на восточном пределе генетического влияния ели европейской // Вестник Оренбургского государственного университета. 2011. №6 (125). С. 125–130.

35. *Ивантер Э.В., Коросов А.В.* Введение в количественную биологию. Петрозаводск, 2000. 318 с.

36. *Иванюкович Г.А.* Статистический анализ экогеологических данных: Практикум решения задач с помощью пакета программ Statistica. СПб.: СПбГУ, 2010. 204 с.

37. *Камалдинов Е.В.* Каноническая дискриминантная модель влияния генотипа отца на некоторые интерьерные показатели его потомков у свиней // Вестник Красноярского государственного аграрного университета. 2012. №1. С. 117–122.

38. *Кендалл М., Стьюарт А.* Многомерный статистический анализ и временные ряды. М.: Наука, 1976. 736 с.

39. *Кирпичников В.С.* Генетика, селекция и гибридизация рыб. Л.: Наука, 1987. 520 с.

40. *Киртбая Е.К.* Биологические основы селекции ягодных культур в условиях Северного Кавказа: дис. … д–ра с.–х. наук. Краснодар, 1989. 543 с.

41. *Клекка У.Р.* Дискриминантный анализ // Факторный, дискриминантный и кластерный анализ. М.: Финансы и статистика, 1989. С. 78–137.

42. Компьютерная биометрика / под ред. В.Н. Носова. М.: Изд-во МГУ, 1990. 232 с.

43. *Кононков П.Ф. [и др.].* Технология выращивания и переработки листовой массы амаранта как сырья для пищевой промышленности. М.: Рос. ин-т дружбы народов, 2008. 195 с.

44. *Корона В.В., Васильев А.Г.* Строение и изменчивость листьев растений: Основы модульной теории. Екатеринбург: Екатеринбург, 2000. 224 с.

45. *Кузнецова А.П. [и др.].* Выделение эффективных источников устойчивости к коккомикозу из форм подвоев мелкокосточковых селекции СКЗНИИСиВ // Плодоводство и ягодоводство России. 2012. Т. 34, №4. С. 439–445.

46. *Кузнецова А.П., Щеглов С.Н.* Системный анализ изменчивости при создании методов ускоренной оценки устойчивости форм *Cerasus* Mill. к коккомикозу // Сборник научных трудов Никитского ботанического сада. 2009. Т. 131. С. 103–107.

47. *Кулаичев А.П.* Методы и средства анализа данных в среде Windows. STADIA 6.0. М.: Информатика и компьютеры, 1996. 257 с.

48. *Купцов А.И.* Изменение корреляций у растений при отборе // ДАН СССР. 1948. Т. 61, № 3. С. 57–66.

49. *Лакин Г.Ф.* Биометрия: учеб. пособие для вузов. М.: Высшая школа, 1990. 352 с.

50. *Любищев А.А.* О количественной оценке сходства // Применение математических методов в биологии. Л.: Изд-во ЛГУ, 1962. С. 152–160.

51. *Любищев А.А.* О приложении математической статистики к практической систематике // Прикладная математика в биологии. М.: Изд-во МГУ, 1979. С. 12–28.

52. *Любищев А.А.* О применении биометрии в систематике // Вестник ЛГУ. 1959. №9. С. 128–136.

53. *Мамаев С.*А., *Попов П.П.* Ель сибирская на Урале (внутривидовая изменчивость и структура популяций). М.: Наука, 1989. 104 с.

54. *Маркелов И.Н., Нижегородцев А.А.* Дискриминация культурных форм примулы многоцветковой (*Primula* × *Polyantha*

Miller, 1768) по геометрическим характеристикам венчика цветка // Вестник Нижегородского университета им. Н.И. Лобачевского. 2014. №1–1 (1). С. 166–171.

55. *Мержаниан А.С.* Селекция в виноградарстве. Ростов н/Д, 1928. 185 с.

56. *Мина М.В., Клевезаль Г.А.* Рост животных. М.: Наука, 1976. 291 с.

57. *Моосмюллер Г., Ребик Н.Н.* Маркетинговые исследования с SPSS: учеб. пособие. М.: ИНФРА-М, 2007. 160 с.

58. *Назаретян В.Г. [и др.]*. Использование методов многомерного статистического анализа в исследовании комплекса бактериологических показателей для выявления оптимального способа лечения дисбиоза у детей с хроническим гастродуоденитом // Кубанский научный медицинский вестник. 2004. № 5–6 (71–72). С.77–80.

59. *Наследов А.Д.* Математические методы психологического исследования. Анализ и интерпретация данных: учеб. пособие. СПб.: Речь, 2004. 392 с.

60. *Неповинных А.Г.* Прогнозирование строения и роста сосняков Красноярской лесостепи: автореф. дис. … канд. с.–х. наук. Красноярск, 2009. 18 с.

61. *Олонова М.В., Хисориев Х.Х.* Систематическая структура комплекса признаков *Poa nemoralis* L. и *P. relaxa* Ovcz. в Гиссаро-Дарвазском флористическом районе Таджикистана // Доклады Академии наук Республики Таджикистан. 2013. Т. 56, №3. С. 243–249.

62. *Петрачук Е.С., Янкова Н.В., Крохалевский В.Р.* Экологическая изменчивость морфометрических признаков леща *Abramis brama* (L., 1758) Обь-Иртышского бассейна // Аграрный вестник Урала. 2013. №1 (107). С. 55–57.

63. *Пинчук В.В., Тихомиров В.Н.* Изменчивость *Epipactis helleborine* и *Epipactis atrorubens* (*Orchidaceae*) при их гибридизации // Вестник БГУ. Сер. 2: Химия. Биология. География. 2010. №1. С. 39–44.

64. *Попов П.П.* Географическая изменчивость формы семенных чешуй ели в Восточной Европе и Западной Сибири // Лесоведение. 1999. №1. С. 68–73.

65. *Попов П.П.* Ель на востоке Европы и в Западной Сибири. Новосибирск: Наука, 1999. 169 с.

66. *Попов П.П.* Классификация деревьев по форме семенных чешуй в промежуточных популяциях ели европейской и сибирской с помощью дискриминантного анализа // Вестник экологии, лесоведения и ландшафтоведения. 2009. №9. С. 77–85.

67. *Попов П.П.* Формирование структуры популяций и популяционно-географическая структура ареала ели европейской и сибирской // Вестник экологии, лесоведения и ландшафтоведения. 2010. №10. С. 116–125.

68. *Попов П.П., Зенкова Е.Л.* Сходство в географической изменчивости ели по фенетическим и генетическим признакам // Вестник экологии, лесоведения и ландшафтоведения. 2009. №10. С. 126–132.

69. *Правдин И.Ф.* Руководство по изучению рыб. М.: Пищевая промышленность, 1966. 375 с.

70. *Правдин Л.Ф.* Ель европейская и ель сибирская в СССР. М.: Наука, 1975. 176 с.

71. *Пузаченко Ю.Г.* Математические методы в экологических и географических исследованиях. М.: Академия, 2004. 416 с.

72. *Сажин Ю.В., Басова В.А.* Многомерные статистические методы: учеб. пособие. М.: Спутник, 2002. 163 с.

73. *Смирнова Н.С. [и др.]*. Влияние динамики фотосинтетических пигментов при созревании подсолнечника, обработанного биопрепаратами перед посевом, на величину урожая и масличность семян // Известия вузов. Пищевая технология. 2008. № 5–6. С. 33–35.

74. *Смирнова О.В*. Ценопопуляции растений (основные понятия и структура). М.: Наука, 1976. 217 с.

75. *Танков Д.А.* Прогноз послепожарного состояния деревьев сосны обыкновенной в низкополнотных древостоях и рединах в условиях степной лесорастительной зоны Оренбургской области // Вестник Оренбургского государственного педагогического университета. 2013. №2. С. 63–66.

76. *Трухачёва Н.В*. Математическая статистика в медико-биологических исследованиях с применением пакета STATISTICA. М.: ГОЭТАР-Медиа, 2012. 384 с.

77. *Тюрин В.В.* Анализ изменчивости комплексов количественных признаков как методология эколого-генетического изучения селекционируемых и естественных популяций рыб: автореф. дис. … д–ра биол. наук. Краснодар, 2010. 46 с.

78. *Тюрин В.В.* Минимизация эффектов факторов среды в селекционной оценке прудовых рыб по продуктивности. Автореф. дис. … канд. биол. наук. М., 1994. 22 с.

79. *Тюрин В.В.* Подход к разработке критерия отбора семей растительноядных рыб по продуктивности с точки зрения многомерной статистики // Международная научно-практическая конференция «Проблемы воспроизводства растительноядных рыб и их роль в аквакультуре» Адлер, 2000. С. 49–51.

80. *Тюрин В.В.* Принцип апостериорной минимизации эффекта модификационной изменчивости на примере оценки продуктивности в семейной селекции белого толстолобика (*Hypophtalmicthis molitrics. Val.*) // Труды Кубанского государственного аграрного университета. 2008. № 5 (14). С. 118–121.

81. *Тюрин В.В., Бурло А.М.* Динамика продуктивности в ряду поколений синтетической селекции местного карпа с использованием интродуцированных пород // Труды Кубанского государственного аграрного университета. 2009. № 1 (16). С. 93–98.

82. *Тюрин В.В., Волчков Ю.А.* Определение минимального числа признаков, разделяющих генетически различные группы рыб в системном морфометрическом анализе // Труды Кубанского государственного аграрного университета. 2009. № 5 (20) С. 132–136.

83. *Тюрин В.В., Волчков Ю.А.* Оценка продуктивности при прудовом выращивании рыб по принципу апостериорной минимизации эффекта факторов среды. Деп. в ВИНИТИ №1973-В93 14.07.1993 11с.

84. *Тюрин В.В. [и др.]*. Информативная селекционная оценка пород карпа по признакам продуктивности при прудовом выращивании // Сборник научных трудов ВНИИПРХ. М., 1992. Вып. 68. С. 72–88.

85. *Тюрин В.В., Морев И.А., Волчков Ю.А.* Дискриминантный анализ в селекционно-генетических исследованиях: практикум. Краснодар: Кубанский гос. ун-т, 2003. 24 с.

86. *Тюрин Ю.Н., Макаров А.А.* Статистический анализ данных на компьютере. М.: ИНФРА-М, 1998. 528 с.

87. *Ушаков М.В., Зиненко А.И.* Подвидовая принадлежность обыкновенной гадюки (*Serpentes*: *Viperidae*) из Воронежской и Липецкой областей // Вестник Тамбовского университета. Сер.:

Естественные и технические науки. 2013. Т.18, №6–1. С. 3093–3097.

88. *Филипченко Ю.А.* Генетика мягких пшениц. М.: Наука, 1979. 312 с.

89. *Филипченко Ю.А.* Изменчивость и методы её изучения. М.; Л., 1929. 275 с.

90. *Хайров И.Х., Кабанов С.В.* Возрастная структура ценопопуляций *Populus tremula* L. в нагорных лесах юга Приволжской возвышенности // Вестник Алтайского государственного аграрного университета. 2011. Т. 84, №10. С. 37–46.

91. *Халафян А.А.* STATISTICA 6. Статистический анализ данных: учебник. М.: Бином-Пресс, 2008. 512 с.

92. *Халафян А.А.* Статистический анализ данных. STATASTICA 6.0. Краснодар: Кубанский гос. ун-т, 2005. 307 с.

93. *Чащин П.В., Чащина О.Е., Захарова Е.Ю.* Использование не дифференцированных по полу выборок в исследованиях структуры байбака (*Marmota bobak* Mull.) // Известия Челябинского научного центра УрО РАН. 2004. №4. С. 152–155.

94. *Чудинова Ю.В.* Оценка продуктивности гибридных и родительских форм льна в условиях Тюменской области // Вестник Томского государственного педагогического университета. 2007. №6. С. 38–42.

95. *Щеглов Н.И.* Сравнительная генетика растений: учеб. пособие. Краснодар: Кубанский гос. ун-т, 2015. 51 с.

96. *Щеглов С.Н.* Изменчивость и методы ее изучения в селекции ягодных культур. Краснодар: Кубанский гос. ун-т, 2013. 307 с.

97. *Щеглов С.Н.* Применение биометрических методов для ускорения селекционного процесса плодовых и ягодных культур. Краснодар: Кубанский гос. ун-т, 2005. 106 с.

98. *Юнкеров В.И., Григорьев С.Г.* Математико-статистическая обработка данных медицинских исследований. СПб.: ВМедА, 2002. 266 с.

99. *Яблоков А.В.* Состояние исследований и некоторые проблемы фенетики популяций // Фенетика популяций. М.: Наука, 1982. С. 3–14.

ПРИЛОЖЕНИЯ

Приложение 1

Данные морфометрического описания семей белого толстолобика

Блок 1

№	l	C	r	o	oP	io	AD	pD	AV	pV	lp	VA
1	339	88	24	13	54	42	176	138	155	74	74	87
2	380	93	28	13	68	45	199	154	170	84	80	95
3	349	85	33	13	55	39	179	138	159	78	73	100
4	356	88	28	13	55	43	183	144	163	82	75	90
5	360	90	27	13	57	43	179	150	166	81	71	94
6	349	88	30	12	55	41	182	142	158	76	76	83
7	349	88	30	11	55	38	173	140	159	73	63	88
8	391	98	32	13	63	43	196	165	177	82	81	94
9	327	85	25	12	53	41	170	135	152	72	68	82
10	335	85	25	13	53	40	174	140	155	73	73	85
11	350	89	28	13	55	39	180	140	162	78	76	92
12	387	93	29	13	57	43	194	158	171	90	72	94
13	345	89	31	13	56	41	180	140	156	75	68	95
14	340	86	31	11	53	39	178	141	151	73	69	84
15	347	91	26	11	53	39	178	142	154	77	72	81
16	377	91	29	12	56	42	191	153	164	82	70	88
17	376	91	28	12	57	41	193	169	159	81	75	89
18	364	90	28	13	57	40	183	154	164	83	72	90
19	333	84	28	11	51	39	172	143	147	74	66	83
20	390	93	28	13	57	42	196	167	171	85	76	95
21	355	89	28	13	55	40	182	152	162	81	72	89
22	378	91	29	12	58	43	194	159	168	80	80	96
23	345	83	25	12	54	39	179	138	161	74	65	87
24	352	88	25	11	53	40	174	152	160	70	75	84
25	342	88	25	11	52	41	170	140	156	72	70	82
26	356	90	27	11	56	40	177	147	164	79	74	89
27	405	102	34	13	59	46	205	159	185	91	79	110
28	379	96	27	12	61	45	196	146	175	85	75	111
29	351	87	27	11	56	39	174	145	158	75	74	89
30	355	90	28	12	56	42	180	146	170	86	75	90
31	395	96	29	12	60	44	199	160	175	83	71	99
32	382	97	30	12	60	44	198	151	178	87	76	97
33	343	85	26	12	51	39	176	135	159	74	68	91
34	343	86	27	11	54	38	173	142	156	74	63	86
35	394	95	32	13	57	46	197	162	181	86	70	94
36	332	88	28	12	53	41	182	120	163	75	72	90
37	388	94	30	12	57	43	199	158	176	89	75	85

№	l	C	r	o	oP	io	AD	pD	AV	pV	lp	VA
38	380	93	28	13	56	41	189	159	168	81	69	93
39	329	82	26	12	49	36	166	137	152	76	66	84
40	351	88	24	12	54	40	178	148	160	76	71	98
41	372	88	27	12	58	38	186	154	167	80	71	89
42	352	85	27	12	50	40	179	147	162	78	75	98
43	343	87	27	12	54	40	175	137	151	73	74	85
44	361	90	36	12	55	39	186	143	166	78	85	81
45	352	90	24	12	54	39	180	150	161	75	75	91
46	345	89	25	13	55	40	174	143	156	73	68	88
47	344	88	26	13	55	36	176	140	156	75	87	80
48	337	87	24	12	52	39	172	138	154	69	65	82
49	350	88	27	13	55	42	181	152	165	81	72	85
50	352	85	25	12	53	40	178	150	160	76	70	86
51	343	87	27	12	55	39	174	134	156	74	68	87
52	345	90	27	13	56	40	179	140	156	75	70	89
53	356	90	26	13	56	42	179	152	160	78	69	88
54	343	87	27	12	54	38	175	142	145	76	68	90
55	346	89	25	12	53	41	181	140	157	70	71	91
56	362	94	28	12	60	42	183	143	167	80	79	92
57	364	90	28	12	55	41	180	156	164	81	75	92
58	418	102	35	12	64	48	212	168	194	93	78	104
59	342	87	24	12	55	40	178	125	156	74	73	87
60	370	90	26	12	55	38	182	155	167	80	72	100
61	340	88	25	12	53	39	174	142	154	72	68	85
62	376	96	30	12	58	49	189	151	175	84	76	93
63	356	90	27	12	57	40	183	145	163	80	73	90
64	354	90	26	12	55	40	180	147	160	75	71	96
65	355	89	26	12	55	40	184	138	165	79	68	95
66	356	90	30	11	55	40	182	139	165	78	72	92
67	412	100	31	13	63	50	209	162	186	87	79	96
68	365	92	29	12	57	40	187	150	170	81	75	99
69	365	90	28	13	58	44	187	150	165	80	72	94
70	365	92	30	11	56	40	185	150	172	81	72	93
71	353	90	26	13	54	39	188	142	160	75	69	95
72	334	86	23	12	52	38	172	128	150	72	70	81
73	352	91	28	12	55	38	180	147	161	74	73	90
74	385	95	26	13	60	45	195	162	173	82	82	98
75	345	87	26	12	55	37	174	139	157	77	70	90
76	360	85	27	12	54	41	182	145	163	80	73	93
77	323	84	25	11	52	40	167	133	152	70	61	83
78	365	92	29	13	56	43	188	149	168	78	73	95

№	l	C	r	o	oP	io	AD	pD	AV	pV	lp	VA
79	368	93	28	11	56	41	187	155	172	84	75	99
80	332	83	27	12	50	39	170	135	150	68	68	82
81	381	95	30	12	58	43	195	153	175	85	78	108
82	385	95	27	13	56	43	196	158	175	85	76	91
83	360	94	26	13	58	42	179	154	160	77	74	98
84	348	90	29	12	55	41	182	143	161	77	66	90
85	340	87	25	11	38	44	172	141	160	76	68	87
86	344	88	25	12	54	41	178	136	161	78	72	88
87	330	80	24	11	49	38	167	140	153	75	62	88

Блок 2

№	IV	H	h	ld	hd	lA	hA	AC	T	P	Семья
1	59	102	43	38	59	55	45	59	41	780	M3
2	63	113	34	40	64	60	45	65	48	1014	N2
3	55	108	40	36	66	44	46	56	45	802	N2
4	57	107	35	38	66	58	45	63	46	858	M2
5	55	107	34	38	62	57	45	65	43	866	M2
6	56	103	38	38	59	60	48	62	45	821	M3
7	56	100	36	36	65	57	45	63	44	788	M2
8	62	119	40	39	71	64	57	75	48	1128	M2
9	50	102	39	35	60	53	45	55	42	688	M3
10	58	100	32	36	68	52	45	61	41	702	M3
11	60	111	36	36	70	55	45	59	46	800	M2
12	60	122	39	43	65	63	49	68	48	1043	M2
13	56	108	38	36	64	56	42	58	42	758	M3
14	51	102	36	34	63	56	45	61	42	698	M3
15	53	106	32	36	62	65	46	61	43	709	M3
16	56	110	38	38	66	57	45	65	49	946	N2
17	56	111	41	38	69	63	46	62	46	970	N2
18	56	103	38	36	63	53	45	66	45	848	N2
19	53	102	36	35	53	54	43	57	40	673	M3
20	60	110	42	40	70	59	46	73	50	1042	N2
21	54	106	37	37	68	56	42	65	43	836	M2
22	56	110	41	36	70	57	48	73	51	988	N2
23	55	105	31	36	60	63	72	60	51	775	M2
24	40	98	32	35	71	55	45	66	40	798	N2
25	54	98	37	35	64	55	44	57	42	688	M3
26	59	103	36	37	70	58	45	60	44	838	M2
27	60	111	39	45	77	61	54	67	52	1179	N2
28	62	116	40	43	70	53	47	56	46	1066	N2

№	IV	H	h	ld	hd	lA	hA	AC	T	P	Семья
29	56	102	35	37	68	55	44	62	41	790	M2
30	60	105	35	38	70	56	47	62	41	846	M2
31	60	112	42	41	68	64	46	72	51	1067	N2
32	56	113	35	40	69	59	47	65	45	1066	M2
33	56	99	36	36	64	53	43	60	44	748	M3
34	50	100	32	35	70	57	46	62	41	753	M2
35	57	108	37	42,1	65	68		70	48	1080	N2
36	55	100	38	35	68	52	45	52	41	782	M3
37	60	109	39	41	70	70	50	66	45	1044	N2
38	50	110	39	42	71	60	49	71	47	985	N2
39	50	93	34	33	59	52	39	62	41	653	N2
40	55	103	34	38	68	58	44	62	43	824	M2
41	56	104	36	38	59	57	45	66	43	870	M2
42	57	104	36	36	69	52	44	66	44	827	M2
43	54,5	100	36	37	63	58	46	60	43	755	M3
44	55	106	38	39	65	56	46	65	44	868	M2
45	58	102	36	37	64	60	47	64	45	820	M2
46	54	99	36	37	61	58	45	61	44	854	N2
47	52	102	32	37	57	56	45	60	42	736	M2
48	54	99	37	35	65	57	44	57	39	720	M3
49	57	106	36	36	72	56	43	62	47	845	M2
50	47	100	38	35	67	56	45	65	44	802	M3
51	54	102	36	35	66	54	45	62	43	744	M3
52	53	100	36	36	65	57	46	62	41	747	M3
53	58	103	36	36	71	54	45	61	45	866	M2
54	55	102	38	37	64	53	43	58	43	780	M2
55	52	99	38	36	68	57	45	60	41	754	M3
56	62	110	38	40	70	61	54	62	48	964	M2
57	54	102	38	38	69	61	48	63	46	885	M2
58	63	115	40	45	81	70	57	77	54	1315	N2
59	55	98	35	38	68	56	50	61	40	755	M3
60	54	104	38	41	72	56	46	68	50	950	N2
61	55	97	35	38	67	55	44	56	42	733	M3
62	56	110	38	42	75	59	48	65	46	1020	M2
63	74	103	35	38	70	57	45	67	45	865	M2
64	55	103	36	40	71	54	47	62	47	850	M2
65	65	103	36	40	68	59	47	59	45	824	M3
66	56	103	36	39	66	58	48	59	45	835	M3
67	58	116	40	44	70	71		71	57	1285	N2
68	55	108	41	38	70	60	50	59	47	990	N2
69	55	105	38	38	70	60	50	68	42	900	M3

№	lV	H	h	ld	hd	lA	hA	AC	T	P	Семья
70	55	104	38	40	70	59	45	65	48	938	M2
71	57	100	38	38	70	57	50	64	42	805	M3
72	56	100	35	38	68	57	48	56	40	735	M3
73	57	105	40	38	68	55	51	65	45	870	N2
74	60	110	38	44	76	61	54	71	47	1105	M2
75	84	99	35	36	66	58	48	62	43	775	M2
76	54	102	36	38	78	60	45	67	43	875	M2
77	52	93	35	34	63	55	43	56	41	665	M3
78	58	105	38	36	66	56	50	62	44	890	M3
79	55	107	40	38	69	62	48	63	46	995	N2
80	42	95	34	36	62	50	42	58	39	670	M2
81	64	109	41	41	73	63	56	62	50	1110	N2
82	62	104	44	41	75	57	51	71	48	1012	N2
83	56	103	35	35	69	55	47	56	45	858	M2
84	55	96	35	36	62	56	48	61	44	785	M3
85	54	95	35	35	57	53	47	62	44	742	M3
86	56	100	38	35	60	54	45	57	43	800	M3
87	49	99	35	33	59	49	40	63	42	714	N2

Примечание. № – порядковый номер строки для объединения первого и второго блока данных. L–P – условные обозначения морфометрических признаков. Семья – переменная, идентифицирующая семейную принадлежность рыб.

Приложение 2

Данные морфометрического описания самцов и самок белого толстолобика

Блок 1

№	L	C	R	O	PO	IO	HH	H	AD	AV
1	189	50	14	9	27	23	55	17	95	83
2	198	49	14	10	27	21	60	19	98	88
3	191	49	15	10	28	24	59	18	98	88
4	195	52	16	11	28	23	60	19	101	90
5	205	52	16	11	28	24	66	20	102	90
6	195	51	16	11	28	21	60	19	99	87
7	195	48	14	10	27	21	57	18	95	86
8	202	49	17	10	27	22	60	18	100	87
9	185	47	14	10	25	20	59	18	99	83
10	197	48	14	11	27	22	61	18	99	89
11	206	52	16	11	30	23	63	19	104	94
12	204	53	16	10	29	23	63	19	106	93
13	207	53	17	11	30	24	66	21	106	92
14	205	53	16	11	29	22	64	20	103	93
15	201	50	15	10	29	23	62	20	102	89
16	197	48	14	10	28	22	62	21	101	89
17	204	51	16	11	30	22	66	19	104	91
18	205	51	16	10	28	23	63	19	103	91
19	195	48	13	9	26	20	59	18	100	88
20	205	52	16	10	28	23	63	20	104	92
21	222	55	19	12	31	26	69	23	111	99
22	209	55	17	10	30	24	64	19	109	97
23	207	52	16	10	29	23	63	19	105	92
24	199	50	16	11	27	22	64	21	102	88
25	209	53	16	10	29	24	64	21	108	95
26	201	52	16	11	29	23	61	21	103	89
27	184	47	14	10	25	21	55	18	94	83
28	207	53	16	11	29	24	64	19	105	95
29	197	50	16	10	27	21	62	20	100	89
30	193	49	18	10	27	21	62	20	96	86
31	209	53	15	9	21	30	62	20	104	96
32	212	55	16	9	21	29	62	21	110	95
33	203	52	15	8	22	29	62	19	105	95
34	202	51	14	8	23	29	60	18	100	94
35	203	52	17	8	23	29	61	19	105	96

№	L	C	R	O	PO	IO	HH	H	AD	AV
36	221	56	15	10	24	32	63	20	113	98
37	204	51	14	8	23	29	58	19	104	94
38	198	52	15	8	22	31	58	19	102	92
39	213	53	15	8	22	30	62	18	108	100
40	203	50	12	9	20	28	57	18	106	90
41	210	52	16	9	22	29	62	18	108	98
42	200	53	15	8	25	29	60	18	107	93
43	218	55	15	9	25	32	63	19	109	101
44	216	56	15	8	25	30	63	19	112	100
45	216	54	15	10	23	30	64	19	107	96
46	211	55	14	9	22	31	62	19	108	103
47	214	55	15	9	22	31	63	19	107	100
48	200	53	14	10	22	30	57	17	107	96
49	206	52	14	9	22	28	61	17	107	93
50	201	51	14	9	22	32	63	20	116	100
51	207	55	16	9	21	31	61	19	110	97
52	219	55	14	10	23	30	63	20	112	97
53	217	54	15	10	22	30	62	18	111	98
54	218	52	15	8	24	28	60	19	107	94
55	212	52	11	8	22	30	62	17	109	91
56	217	54	14	9	22	29	64	20	108	96
57	216	55	15	9	24	31	64	21	111	101
58	218	52	12	8	23	29	65	22	112	98
59	225	56	15	7	25	31	67	22	114	101
60	220	54	14	9	22	30	63	18	114	102

Блок 2

№	PV	LP	VA	LV	LD	HD	LA	HA	AC	POL
1	38	41	49	31	18	41	33	25	33	2
2	43	41	52	32	19	43	33	27	34	2
3	42	40	49	30	18	39	32	25	32	2
4	43	43	51	34	19	42	34	27	33	2
5	45	43	55	35	19	45	33	27	38	2
6	42	43	51	30	19	42	32	26	36	2
7	42	42	48	30	18	42	31	25	34	2
8	43	42	53	34	18	42	33	28	38	2
9	39	40	47	30	18	40	29	24	32	2
10	43	42	52	32	19	40	32	25	34	2
11	44	43	54	33	20	43	32	25	38	2
12	44	44	49	37	20	43	34	28	36	2

№	PV	LP	VA	LV	LD	HD	LA	HA	AC	POL
13	44	42	55	35	19	42	33	28	36	2
14	44	44	53	32	20	44	35	26	37	2
15	45	42	52	32	20	39	33	25	37	2
16	44	43	54	34	19	44	31	26	36	2
17	45	44	54	32	19	43	35	25	36	2
18	46	45	53	32	20	42	34	25	40	2
19	42	40	47	31	19	41	31	25	36	2
20	45	43	55	33	21	44	33	27	38	2
21	49	48	56	37	22	47	36	29	37	2
22	47	44	58	33	21	43	33	25	37	2
23	46	41	54	33	19	42	30	25	39	2
24	45	43	53	32	20	41	31	25	35	2
25	47	45	58	36	21	45	35	27	38	2
26	43	43	51	33	20	41	33	27	35	2
27	39	39	46	30	18	40	29	24	35	2
28	45	42	58	34	20	42	30	25	37	2
29	43	42	51	32	20	41	32	28	35	2
30	42	43	51	33	20	41	31	24	33	2
31	45	40	53	34	22	45	34	29	37	3
32	45	40	54	32	20	41	30	26	37	3
33	41	39	45	32	20	47	33	29	38	3
34	46	40	53	30	19	46	32	28	38	3
35	45	41	54	33	21	47	37	27	35	3
36	46	43	57	34	23	38	36	30	40	3
37	42	40	50	32	21	47	30	28	34	3
38	40	40	51	33	20	45	30	28	33	3
39	44	40	58	34	19	43	32	28	40	3
40	41	39	56	34	20	40	30	28	35	3
41	44	40	54	34	20	44	33	29	37	3
42	44	40	49	31	20	42	30	28	32	3
43	46	39	56	35	20	46	30	29	37	3
44	47	38	54	32	22	47	30	27	36	3
45	46	39	57	33	23	42	34	32	38	3
46	47	42	54	35	20	44	33	29	35	3
47	43	40	56	32	21	44	33	29	38	3
48	43	39	49	32	19	43	30	28	33	3
49	43	41	52	33	20	39	33	29	40	3
50	45	44	53	34	23	43	37	30	41	3
51	42	41	53	32	21	43	30	28	38	3
52	45	42	57	35	22	42	35	30	42	3
53	44	41	55	34	24	43	34	30	37	3

№	PV	LP	VA	LV	LD	HD	LA	HA	AC	POL
54	44	42	48	33	21	43	33	28	36	3
55	42	42	51	34	21	45	33	29	35	3
56	44	41	53	32	24	45	35	29	37	3
57	45	40	51	32	21	47	36	28	39	3
58	45	39	55	32	22	42	35	28	37	3
59	46	45	57	33	22	41	34	29	39	3
60	45	44	57	34	21	44	38	29	35	3

П р и м е ч а н и е. № – порядковый номер строки для объединения первого и второго блока данных. L–AC – условные обозначения морфометрических признаков; Р – переменная, указывающая половую принадлежность рыб (2 – самка, 3 – самец).

Приложение 3

Данные по оценке признаков – компонент продуктивности семей белого толстолобика

pp	pr	mk	m0	ft	Пруд	Семья
915	346	640	115	97,5	20	А2
1000	353	621	116	50,0	21	А2
1790	759	595	117	72,5	22	А2
1648	462	600	116	61,7	23	А2
915	304	502	40	60,0	20	С2
1000	316	492	41	54,0	21	С2
1790	531	457	42	40,0	22	С2
1648	335	475	41	57,9	23	С2
915	329	621	121	56,7	20	С3
1000	304	556	122	100	21	С3
1790	539	545	123	57,5	22	С3
1648	350	576	122	45,8	23	С3

Примечание. Обозначения признаков: pp – плотность посадки; число особей, высаженных в пруд в пересчёте на единицу площади; pr – одна из результирующих характеристик – выход рыбы с единицы площади (кг/га); m0 – средняя масса рыб при посадке в пруд; ft – процент выживаемости; Семья – переменная, идентифицирующая семьи.

Оглавление

Введение……………………………………………………….. 3

1. Теоретические основы дискриминантного анализа…………………………….. 5
2. Реализация дискриминантного анализа в STATISTICA…………………………………… 14
3. Практическое использование дискриминантного анализа……………………………... 34
   - 3.1. Разделение двух групп……………………………… 34
     - 3.1.1. Прижизненная идентификация пола у белого толстолобика……………………. 34
     - 3.1.2. Идентификация форм устойчивых к коккомикозу представителей рода Cerasus Mill. на основании анализа комплекса биохимических показателей……… 38
   - 3.2. Разделение нескольких групп……………………….. 41
     - 3.2.1. Идентификация групповой принадлежности объектов в случае нескольких групп (на примере четырёх пород форели)…………. 41
     - 3.2.2. Сравнительная оценка различных популяций облепихи…………………………… 50
   - 3.3. Определение координат идеального объекта в пространстве дискриминантных функций как метод биологической интерпретации расстояний Махаланобиса……………………….. 53
     - 3.3.1. Оценка продуктивности семей белого толстолобика по принципу апостериорной минимизации средовой компоненты изменчивости……………………………… 54
     - 3.3.2. Динамика продуктивности в ряду поколений синтетической селекции местного карпа с использованием интродуцированных пород.. 61
   - 3.4. Использование дискриминантного анализа в биологии…………………………………………… 68

Заключение…………………………………………………... 106

Библиографический список………………………………… 107

Приложения………………………………………………… 115

Научное издание

Т ю р и н Владислав Викторович
Щ е г л о в Сергей Николаевич

ДИСКРИМИНАНТНЫЙ АНАЛИЗ В БИОЛОГИИ

Монография

Подписано в печать 20.04.2015. Формат 60×84 $^{1}/_{16}$.
Гарнитура «Таймс». Печать офсетная. Уч.-изд. л. 8,2.
Тираж 500 экз. Заказ № 15126.

Кубанский государственный университет
350040, г. Краснодар, ул. Ставропольская, 149.

Типография ООО «Просвещение-Юг»
350080, г. Краснодар, ул. Бородинская, 160/5.
Отпечатано с оригинал-макета заказчика.